

Lösungsblatt 1

Signalverarbeitung

Aufgabe 1: Faltung

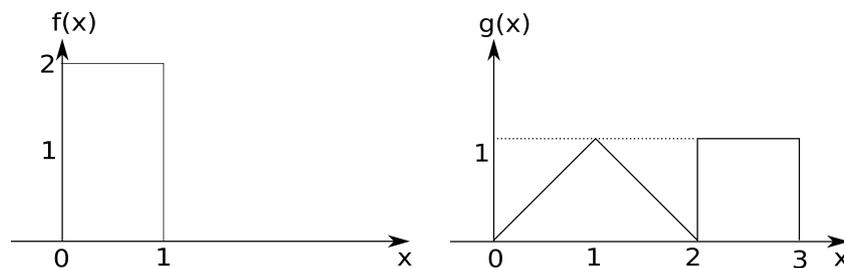


Abbildung 1: Faltung

- a) Bestimmen Sie die Faltung der Funktion $f(x)$ mit sich selbst ($f(x) * f(x)$) rechnerisch.

Lösung:

Rechnung:

$$f(x) * f(x) = \int_{-\infty}^{\infty} f(t)f(x-t)dt$$

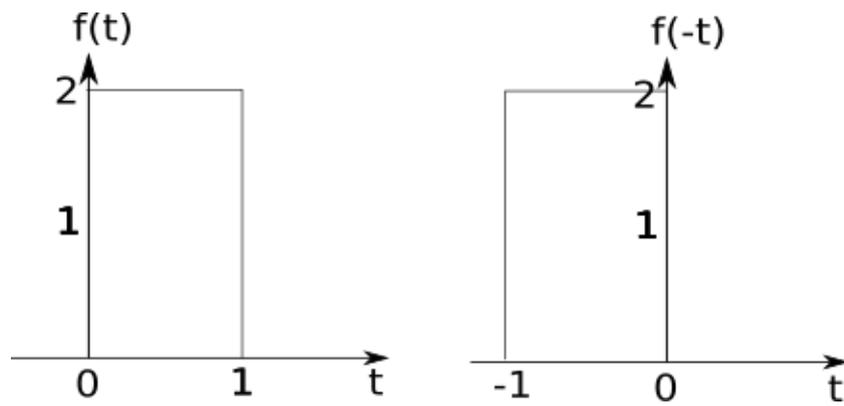
(1)

Gemäß der Definition der Faltung erhält man die Faltung von $f(x)$ mit sich selbst durch die Integration von $f(t)$ mit der gespiegelten Funktion $f(-t)$. Bildlich gesprochen wird für die Integration $f(-t)$ von links über $f(t)$ geschoben.

Fallunterscheidung:

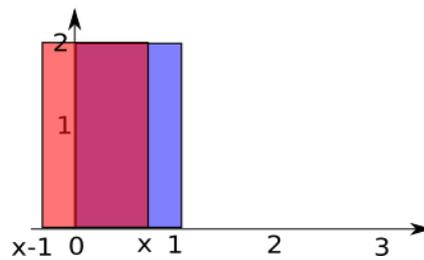
Fall 1:

$$x < 0 \rightarrow \int_{-\infty}^{\infty} f(t) \cdot f(x-t)dt = 0$$



$$x > 2 \rightarrow \int_{-\infty}^{\infty} f(t) \cdot f(x-t) dt = 0$$

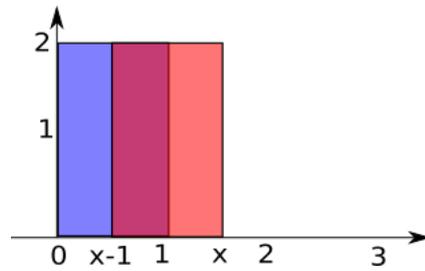
Fall 2:



Wenn die gespiegelte Funktion $f(-x)$ (rotes Rechteck) über $f(x)$ (blaues Rechteck) geschoben wird, entsteht ein Bereich der Überlappung (lila Rechteck).

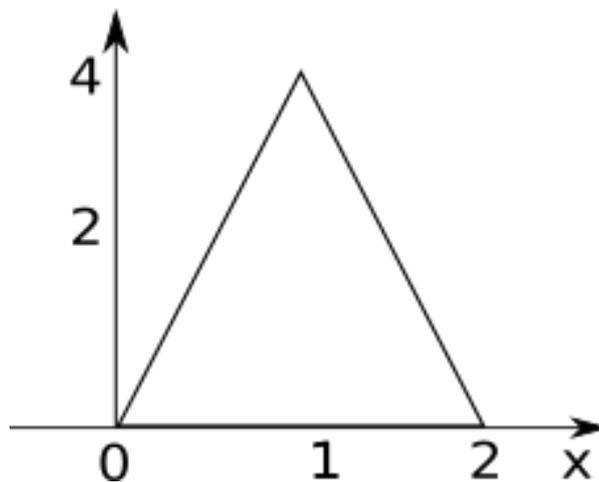
$$0 \leq x < 1 \rightarrow \int_{-\infty}^{\infty} f(t) \cdot f(x-t) dt = \int_0^x 4 dt = 4x$$

Fall 3:



$$\begin{aligned}
 1 \leq x \leq 2 &\rightarrow \int_{-\infty}^{\infty} f(t) \cdot f(x-t) dt \\
 &= \int_{x-1}^1 4 dt = -4x + 8
 \end{aligned}$$

Das Faltungsergebnis ist demnach:



b) Bestimmen Sie $f(x) * g(x)$ grafisch.

Lösung:

Rechnung:

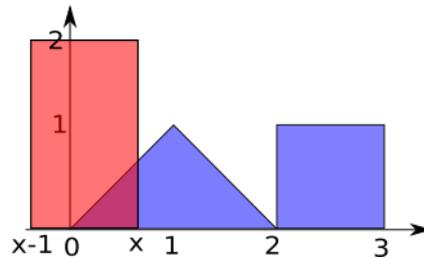
$$f(x) * g(x) = \int_{-\infty}^{\infty} g(t) \cdot f(x-t) dt \quad (2)$$

Fallunterscheidung:

Fall 1:

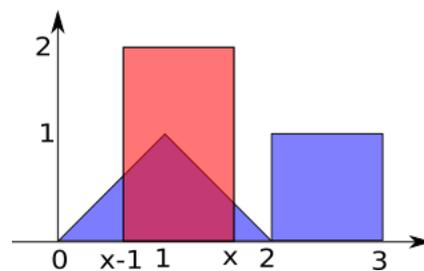
$$x < 0 \rightarrow \int_{-\infty}^{\infty} g(t) \cdot f(x-t) dt = 0$$

Fall 2:



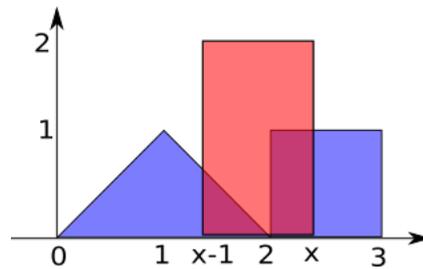
$$\begin{aligned} 0 \leq x < 1 &\rightarrow \int_{-\infty}^{\infty} g(t) \cdot f(x-t) dt \\ &= \int_0^x 2t \cdot dt = \frac{x^2}{2} \cdot 2 \end{aligned}$$

Fall 3:



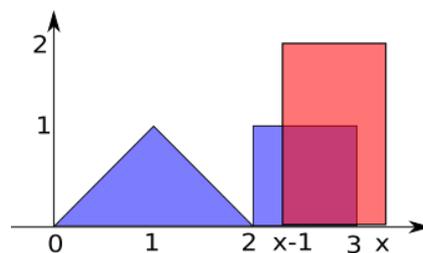
$$\begin{aligned} 1 \leq x < 2 &\rightarrow \int_{-\infty}^{\infty} g(t) \cdot f(x-t) dt \\ &= \int_{x-1}^1 2t \cdot dt + \int_1^x (-2t + 4) \cdot dt \\ &= [t^2]_{x-1}^1 + [-t^2 + 4t]_1^x = -2x^2 + 6x - 3 \end{aligned}$$

Fall 4:



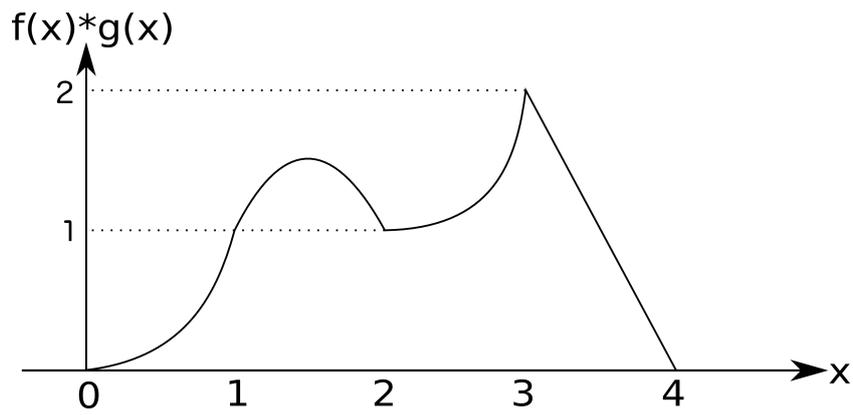
$$\begin{aligned}
 2 \leq x < 3 &\rightarrow \int_{-\infty}^{\infty} g(t) \cdot f(x-t) dt \\
 &= \int_{x-1}^2 (-2t+4) \cdot dt + \int_2^x 2 \cdot dt \\
 &= [-t^2 + 4t]_{x-1}^2 + [2t]_2^x = x^2 - 4x + 5
 \end{aligned}$$

Fall 5:



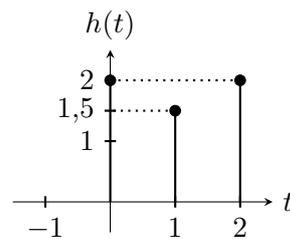
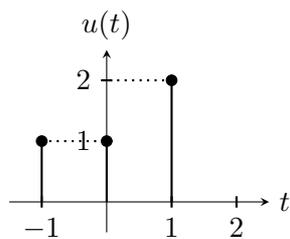
$$\begin{aligned}
 3 \leq x &\rightarrow \int_{-\infty}^{\infty} g(t) \cdot f(x-t) dt \\
 &= \int_{x-1}^3 2 \cdot dt \\
 &= [2t]_{x-1}^3 = -2x + 8
 \end{aligned}$$

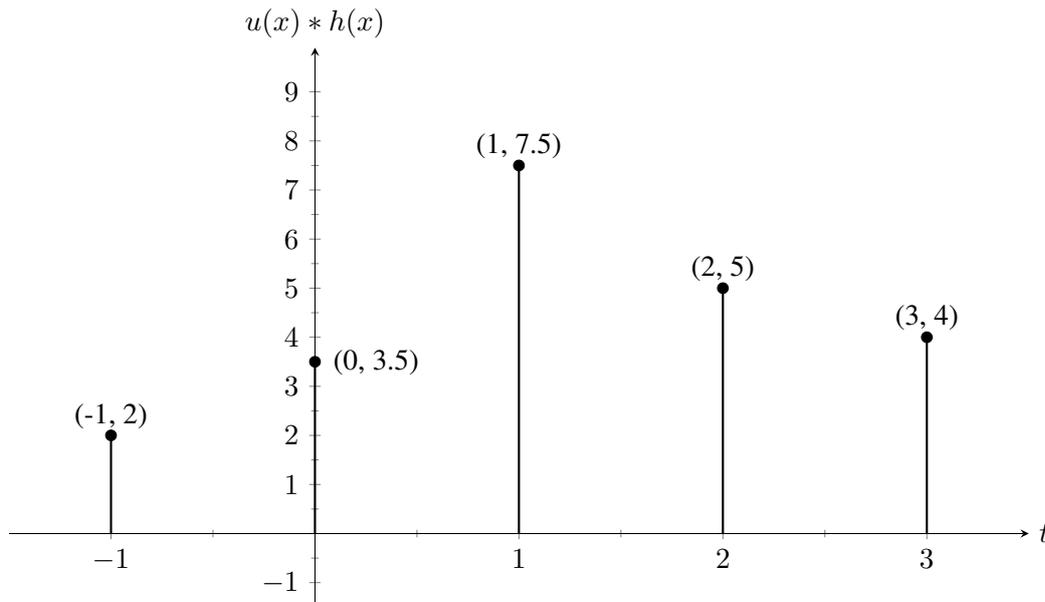
Ausgehend von den vorangegangenen Berechnungen kann das Faltungsergebnis wie folgt visualisiert werden:



c) Bestimmen Sie die diskrete Faltung $u(t) * h(t)$ grafisch.

$$u(t) = \begin{cases} 1 & \text{für } t = -1 \\ 1 & \text{für } t = 0 \\ 2 & \text{für } t = 1 \\ 0 & \text{für sonst} \end{cases}, \quad h(t) = \begin{cases} 2 & \text{für } t = 0 \\ 1,5 & \text{für } t = 1 \\ 2 & \text{für } t = 2 \\ 0 & \text{für sonst} \end{cases}$$





Onlinefrage Nr. 1: Was ist der größte Wert des Ergebnisses der diskreten Faltung $u(x) * h(x)$?

- i) 4,5
- ii) 5,0
- iii) 6,5
- iv) 7,5
- v) 9,0

Lösung:

Der größte Wert des Ergebnisses der Faltung $u(x) * h(x)$ ist 7,5, siehe Lösung der Aufgabe 1c).

Aufgabe 2: Digitalisierung von Signalen

Gegeben sei das Signal $f(t) = \sin(2\pi 200t) + 4 \cos(2\pi 300t)$

Hinweise:

- Zur Vereinfachung sei in der Zeichnung die Höhe von $\delta = 1$
- Zur Fouriertransformation existieren diverse Konventionen! Lassen Sie sich davon nicht verwirren und verwenden Sie zur Lösung dieser Aufgabe die folgenden Informationen:

Funktion	Fouriertransformierte
$\sin(2\pi f_0 \cdot t)$	$\frac{i}{2}(\delta(f + f_0) - \delta(f - f_0))$
$\cos(2\pi f_0 \cdot t)$	$\frac{1}{2}(\delta(f + f_0) + \delta(f - f_0))$
$\sum_{k=-\infty}^{+\infty} \delta(t - kT)$	$\frac{1}{T} \sum_{k=-\infty}^{+\infty} \delta(f - k/T)$

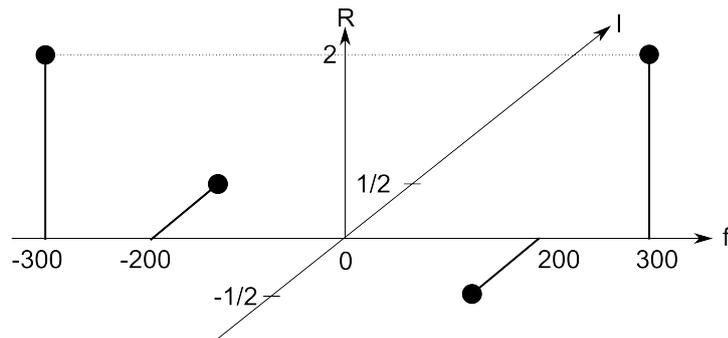
a) Zeichnen Sie das komplexe Spektrum der mit 500 Hz abgetasteten Funktion $f(t)$, t in Sekunden.

Lösung:

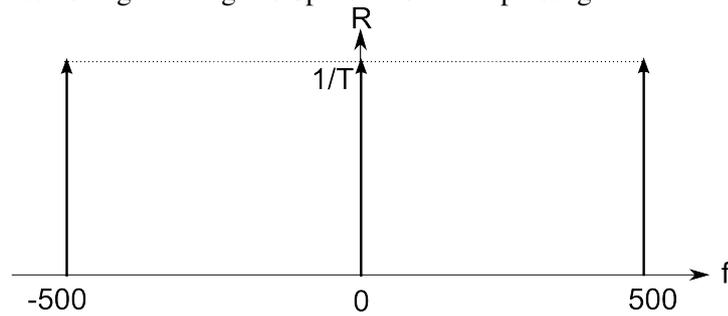
Die Fouriertransformation der Funktion $f(t)$ ist:

$$\begin{aligned}
 F(f) &= \frac{i}{2} \cdot (\delta(f + 200) - \delta(f - 200)) + 4 \cdot \frac{1}{2} \cdot (\delta(f + 300) + \delta(f - 300)) \\
 &= \frac{i}{2} \cdot \delta(f + 200) - \frac{i}{2} \cdot \delta(f - 200) + 2 \cdot \delta(f + 300) + 2 \cdot \delta(f - 300)
 \end{aligned}$$

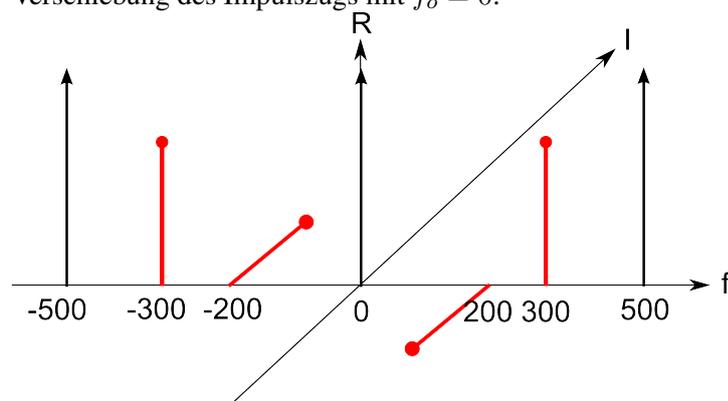
(3)



Abtastung: Faltung des Spektrums mit Impulszug

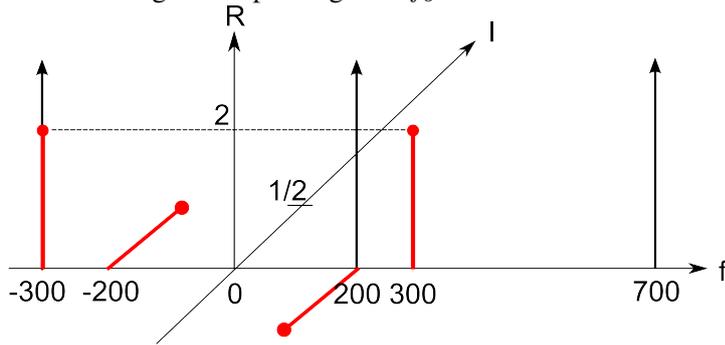


Verschiebung des Impulszugs mit $f_o = 0$:



$$F(0) * \delta(0) = 0$$

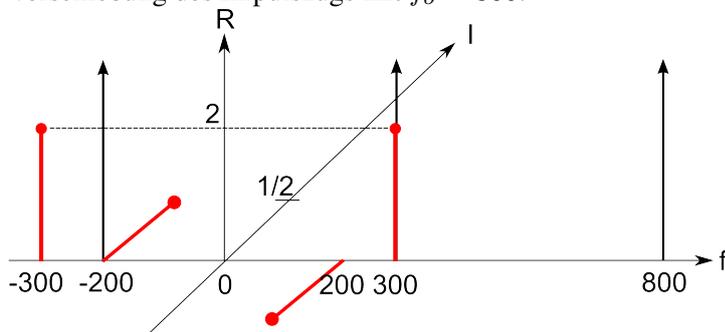
Verschiebung des Impulszugs mit $f_o = 200$:



$$(F * \delta)(f_o = 200) = 2 \cdot \frac{1}{T} - \frac{i}{2} \cdot \frac{1}{T} = \frac{2}{T} - \frac{i}{2T}$$

$$(F * \delta)(f_o = -300) = (F * \delta)(f_o = 200) = (F * \delta)(f_o = 700)$$

Verschiebung des Impulszugs mit $f_o = 300$:



$$(F * \delta)(f_o = 300) = 2 \cdot \frac{1}{T} + \frac{i}{2} \cdot \frac{1}{T} = \frac{2}{T} + \frac{i}{2T}$$

$$(F * \delta)(f_o = -200) = (F * \delta)(f_o = 300) = (F * \delta)(f_o = 800)$$

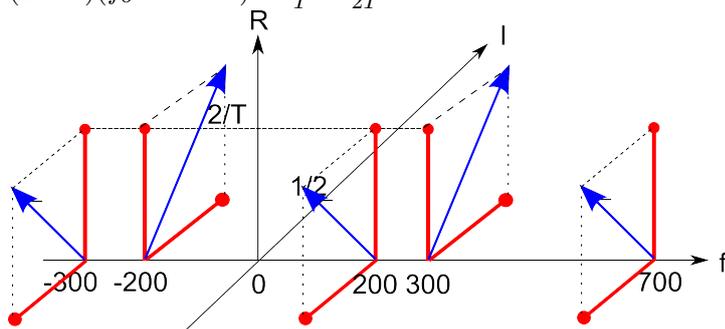
Komplexes Spektrum mit einer Samplefrequenz von 500Hz:

$$(F * \delta)(f_o = 200) = \frac{2}{T} - \frac{i}{2T}$$

$$(F * \delta)(f_o = -200) = \frac{2}{T} + \frac{i}{2T}$$

$$(F * \delta)(f_o = 300) = \frac{2}{T} + \frac{i}{2T}$$

$$(F * \delta)(f_o = -300) = \frac{2}{T} - \frac{i}{2T}$$

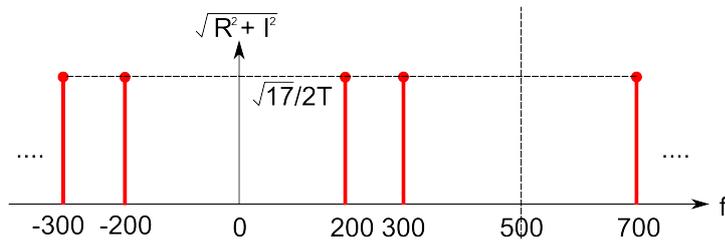


Berechnung des Betragsspektrums nach Abtastung mit 500Hz:

$$|(F * \delta)(f_o = 200)| = \sqrt{\left(\frac{2}{T}\right)^2 + \left(\frac{-1}{2T}\right)^2} = \frac{\sqrt{17}}{2T}$$

$$|(F * \delta)(f_o = 300)| = \sqrt{\left(\frac{2}{T}\right)^2 + \left(\frac{1}{2T}\right)^2} = \frac{\sqrt{17}}{2T}$$

...



b) Welches Phänomen tritt auf, wie kommt es dazu und wie kann verhindert werden, dass es auftritt?

Lösung:

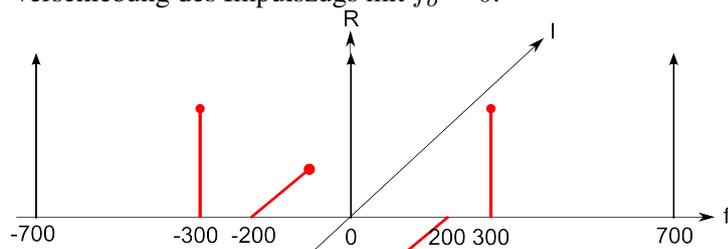
- Die gegebene Abtastfrequenz von 500Hz wurde gemäß des Abtasttheorems zu klein gewählt. Daher tritt Aliasing auf, d.h. wiederholte Spektren überlagern sich.
- Im gegebenen Fall treten Frequenzen auf, die im Originalsignal nicht vorhanden sind, und im Originalsignal vorhandene Frequenzen verschwinden.
- Bei einer D/A-Wandlung kann das Originalsignal nicht mehr aus den abgetasteten Werten rekonstruiert werden.
- Um Aliasing zu verhindern, muss als Abtastfrequenz eine Frequenz gewählt werden, die echt größer ist als das 2-fache der höchsten Frequenz im Signal, also $> 2 \cdot 300$ Hz.

c) Zeichnen Sie das Betragsspektrum der mit 700 Hz abgetasteten Funktion.

Lösung:

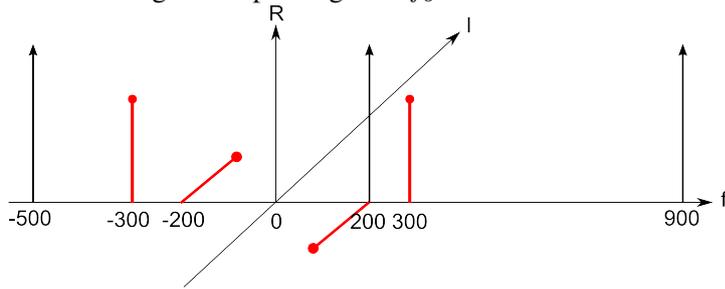
Faltung des Spektrums vor der Abtastung mit dem Impulszug mit Frequenz 700Hz:

Verschiebung des Impulszugs mit $f_o = 0$:



$$(F * \delta)(f_o = 0) = 0$$

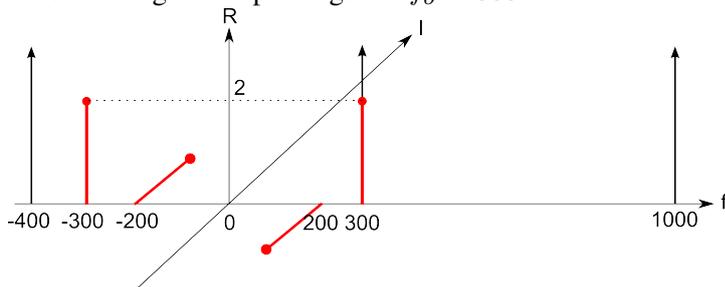
Verschiebung des Impulszugs mit $f_o = 200$:



$$(F * \delta)(f_o = 200) = \frac{-i}{2} \cdot \frac{1}{T} = -\frac{i}{2T}$$

$$(F * \delta)(f_o = -200) = (F * \delta)(f_o = 500) = \frac{i}{2} \cdot \frac{1}{T} = \frac{i}{2T}$$

Verschiebung des Impulszugs mit $f_o = 300$:



$$(F * \delta)(f_o = 300) = \frac{1}{T} \cdot 2 = \frac{2}{T}$$

$$(F * \delta)(f_o = -300) = (F * \delta)(f_o = 400) = \frac{1}{T} \cdot 2 = \frac{2}{T}$$

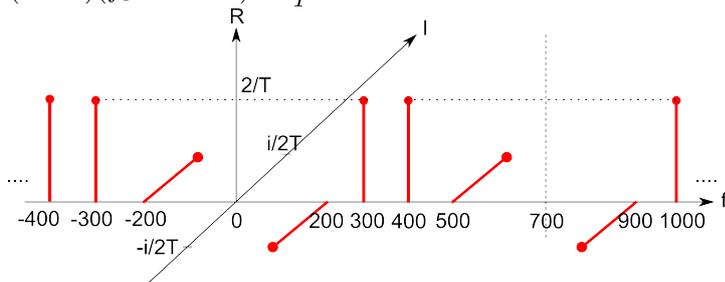
Das komplexe Spektrum nimmt nach der Abtastung folgende Werte an:

$$(F * \delta)(f_o = 200) = -\frac{i}{2T}$$

$$(F * \delta)(f_o = 300) = \frac{2}{T}$$

$$(F * \delta)(f_o = -200) = \frac{i}{2T}$$

$$(F * \delta)(f_o = -300) = \frac{2}{T}$$

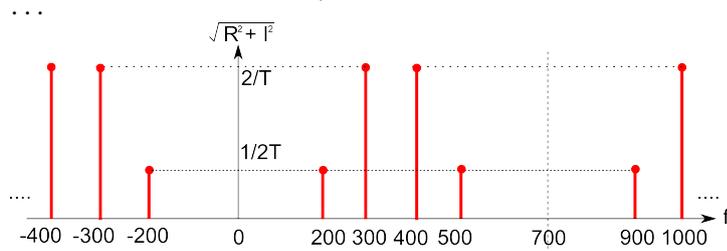


Berechnung des Betragsspektrums nach Abtastung mit 700Hz:

$$|(F * \delta)(f_o = 200)| = \sqrt{0 + \left(\frac{1}{2T}\right)^2} = \frac{1}{2T}$$

$$|(F * \delta)(f_o = 300)| = \sqrt{\left(\frac{2}{T}\right)^2 + 0} = \frac{2}{T}$$

$$|(F * \delta)(f_o = -200)| = \sqrt{0 + \left(\frac{1}{2T}\right)^2} = \frac{1}{2T}$$



Bei einer Abtastfrequenz von 700Hz wird das Abtasttheorem befolgt und es kommt nicht zu Aliasing.

Onlinefrage Nr. 2: Beim Rekonstruieren des Signals aus dem abgetasteten Spektrum aus Aufgabe 2a) entsteht ein Signal mit verfälschter Frequenz. Welche der folgenden Frequenzen sind verfälscht? Beim Ausfüllen der Ilias Onlinefrage geben Sie richtig für die verfälschten Frequenzen an und falsch für die nicht verfälschten Frequenzen.

- i) 0 Hz
- ii) 100 Hz
- iii) 200 Hz
- iv) 300 Hz

Lösung:

Die Frequenzen 200 Hz und 300 Hz sind verfälscht, siehe Zeichnung in 2a).

Aufgabe 3: Filtern

- a) Gegeben sei die Funktion $v(t) = \pi \cdot \delta(t) - \frac{\sin(6\omega_0 t) - \sin(3\omega_0 t)}{t}$ mit der Zeitvariablen t und $\omega_0 > 0$. Die Multiplikation ihrer Fouriertransformierten $V(\omega)$ mit der Fouriertransformierten eines Signals stellt einen Filter dar. Um welchen Filter handelt es sich dabei? Berechnen Sie für diese Aufgabe $V(\omega)$ und vereinfachen Sie in geeigneter Weise.

Hinweis:

Die Funktion $f(t) = \frac{\sin(\pi t)}{\pi t}$ entspricht der Funktion $f(t) = \text{sinc}(t)$. Benutzen Sie als Fouriertransformierte von $g(t) = \text{sinc}(at)$ folgende Funktion: $G(\omega) = \frac{1}{|a|} \text{rect}\left(\frac{\omega}{2\pi a}\right)$ mit $\text{rect}(t) = \begin{cases} 0, & |t| > \frac{1}{2} \\ 1, & |t| \leq \frac{1}{2} \end{cases}$

Lösung:

Umformung:

$$v(t) = \pi\delta(t) - \frac{\sin(\omega_0 t) - \sin(6\omega_0 t)}{t} \quad (4)$$

$$= \pi\delta(t) - \left(\frac{6\omega_0 \sin(6\omega_0 t)}{6\omega_0 t} - \frac{3\omega_0 \sin(3\omega_0 t)}{3\omega_0 t} \right) \quad (5)$$

$$= \pi\delta(t) - \left(6\omega_0 \frac{\sin(\pi \frac{6\omega_0 t}{\pi})}{\pi \frac{6\omega_0 t}{\pi}} - 3\omega_0 \frac{\sin(\pi \frac{3\omega_0 t}{\pi})}{\pi \frac{3\omega_0 t}{\pi}} \right) \quad (6)$$

$$= \pi\delta(t) - \left(6\omega_0 \operatorname{sinc}\left(\frac{6\omega_0 t}{\pi}\right) - 3\omega_0 \operatorname{sinc}\left(\frac{3\omega_0 t}{\pi}\right) \right)$$

Fouriertransformation: Laut Vorlesung ist die Fouriertransformierte von $f(x) = \delta(x)$ die Funktion $F(\omega) = 1$.

$$V(\omega) = F(v(t)) = F(\pi\delta(t) - (6\omega_0 \operatorname{sinc}(\frac{6\omega_0 t}{\pi}) - 3\omega_0 \operatorname{sinc}(\frac{3\omega_0 t}{\pi})))$$

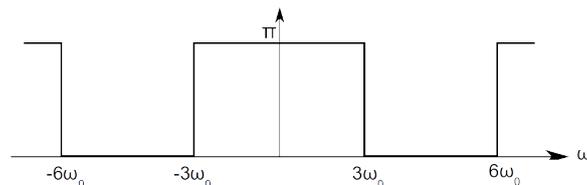
$$\begin{aligned} V(\omega) &= \pi \cdot 1 - \left(6\omega_0 \frac{1}{|\frac{6\omega_0}{\pi}|} \operatorname{rect}\left(\frac{\omega}{2\pi \frac{6\omega_0}{\pi}}\right) - 3\omega_0 \frac{1}{|\frac{3\omega_0}{\pi}|} \operatorname{rect}\left(\frac{\omega}{2\pi \frac{3\omega_0}{\pi}}\right) \right) \\ &= \pi - \left(6\omega_0 \frac{\pi}{6\omega_0} \operatorname{rect}\left(\frac{\omega}{12\omega_0}\right) - 3\omega_0 \frac{\pi}{3\omega_0} \operatorname{rect}\left(\frac{\omega}{6\omega_0}\right) \right) \end{aligned} \quad (7)$$

$$= \pi - \left(\pi \operatorname{rect}\left(\frac{\omega}{12\omega_0}\right) - \pi \operatorname{rect}\left(\frac{\omega}{6\omega_0}\right) \right) \quad (8)$$

$$= \pi - \pi \operatorname{rect}\left(\frac{\omega}{12\omega_0}\right) + \pi \operatorname{rect}\left(\frac{\omega}{6\omega_0}\right) \quad (9)$$

$$(10)$$

- $\operatorname{rect}(t)$ ist definiert als $\begin{cases} 0, & |t| > \frac{1}{2} \\ 1, & |t| \leq \frac{1}{2} \end{cases}$
- $\operatorname{rect}(\frac{\omega}{6\omega_0})$ ist 1 nur dann wenn $|\omega| \leq 3\omega_0$
- $\operatorname{rect}(\frac{\omega}{12\omega_0})$ ist 1 nur dann wenn $|\omega| \leq 6\omega_0$



Onlinefrage Nr. 3: Um welches Filter handelt es sich in Aufgabe 3a)?

- Hochpass
- Bandsperre
- Tiefpass
- nichts davon

Lösung:

Es handelt sich um eine Bandsperre (ii; siehe Lösung 3a).

Aufgabe 4: Diskrete Fouriertransformation, Sampling

Die Frequenzauflösung der DFT hängt von der zeitlichen Auflösung ab. Je feiner die zeitliche Auflösung ist, desto gröber ist die Frequenzauflösung und umgekehrt. Berechnet man beispielsweise die DFT über einem gesamten Musikstück, so erhält man detaillierte Informationen über die auftretenden Frequenzen, dies jedoch nur kumuliert über das gesamte Musikstück hinweg. Berechnet man die DFT auf Basis von beispielsweise 100 Werten, so erhält man als Ergebnis ebenfalls 100 Werte. Da das Resultat der DFT symmetrisch ist, trägt nur die Hälfte der Werte Informationen. Somit verteilt sich die Frequenzinformation auf 50 Werte. Betrachten wir zur Verdeutlichung ein Beispiel mit unterschiedlichen Samplingraten, berechnen dabei die DFT aber stets auf Basis von 100 Samples:

Samplingrate	Grenzfrequenz	Frequenzauflösung	Zeitauflösung
1000 Hz	500 Hz	10 Hz	0.1 s
100 Hz	50 Hz	1 Hz	1 s
10 Hz	5 Hz	0.1 Hz	10 s

Es wird deutlich, dass die Frequenzauflösung in einem reziproken Verhältnis zur zeitlichen Auflösung steht.

- a) Berechnen Sie die Anzahl N der Samples, die benötigt werden, damit eine DFT eine Frequenzauflösung von 5 Hz erreichen kann, wenn ein Signal mit 48 kHz abgetastet wurde.

Lösung:

Bei einer Abtastung mit 48 kHz beträgt die Grenzfrequenz 24 kHz. Bei einer gewünschten Frequenzauflösung von 5 Hz benötigen wir somit 4800 informationstragende Werte. Dies entspricht einer DFT mit 9600 Werten. Folglich benötigen wir $N = 9600$ Samples zur Berechnung der DFT.

Onlinefrage Nr. 4: Welche zeitliche Auflösung wird benötigt, damit eine DFT eine Frequenzauflösung von 5 Hz erreichen kann, wenn ein Signal mit 21 kHz abgetastet wurde?

- i) 0,048 s
- ii) 0,2 s
- iii) 1 s
- iv) 5 s

Lösung:

Aus Teilaufgabe a) wissen wir, dass wir 9600 Samples benötigen. Bei einer Abtastrate von 48 kHz entspricht dies einer zeitlichen Auflösung von 0,2 s.

Lösungsblatt 2

Maschinelles Lernen und Klassifikation

Aufgabe 1: Zufallsexperiment

- a) Nehmen Sie eine Stoppuhr und lassen Sie eine andere Person diese starten. Versuchen Sie, ohne die Stoppuhr zu sehen, nach Ablauf von 5 Sekunden “Stopp” zu sagen, woraufhin die andere Person die Uhr stoppt und den darauf angezeigten Wert notiert. Wiederholen Sie den Versuch n mal, ohne die bereits notierten Werte zu betrachten. Wählen Sie einen sinnvollen Wert für n , um einen Naiven Bayes-Klassifikator zu trainieren.
- b) Nehmen Sie an, das Experiment in Teilaufgabe a) unterliegt einem Zufallsprozess, der mit einer Gauß’schen Normalverteilung beschrieben werden kann. Somit lässt sich für jede Person, die den beschriebenen Versuch durchführt, eine spezifische Wahrscheinlichkeitsdichtefunktion beschreiben. Schätzen Sie die Parameter Ihrer persönlichen Verteilung mit den gemessenen Ergebnissen mit der *Maximum Likelihood*-Methode.
- c) Die in Teilaufgabe b) beschriebene Verteilungsfunktion soll die Wahrscheinlichkeitsdichte für eine Klasse ω_1 beschreiben. Nehmen Sie an, eine Kommilitonin hat einen ähnlichen Versuch durchgeführt. Aus ihren Messungen ergibt sich eine Klasse ω_2 , die ebenfalls einer Normalverteilung unterliegt mit dem Erwartungswert $\mu_2 = 5.1$ und der Varianz $\sigma_2^2 = 0.6$.

Führen Sie einen einzelnen weiteren Versuch wie beim Zufallsexperiment in Teilaufgabe a) durch und versuchen Sie, mit einem Naiven Bayes-Klassifikator zu ermitteln, ob das Ergebnis zu Ihnen selbst oder Ihrer Kommilitonin passt (ω_1 oder ω_2). Sie haben dabei die Handschrift analysiert, mit der das Ergebnis notiert wurde; Sie vermuten deshalb von vornherein, dass es mit einer Wahrscheinlichkeit von 0.4 zu ω_1 gehört und mit einer Wahrscheinlichkeit von 0.6 zu ω_2 .

Lösung:

- a) Für $n = 10$ Zufallszahlen: 2,5; 5,2; 4,5; 6,1; 1,6; 3,6; 3,5; 7,0; 2,7; 6,3
- b) Erwartungswert und Varianz:

$$\begin{aligned}\mu &= \frac{1}{n} \sum_{i=1}^n x_i \\ &= \frac{1}{10} \cdot (2,5 + 5,2 + 4,5 + 6,1 + 1,6 + 3,6 + 3,5 + 7,0 + 2,7 + 6,3) \\ &= 4,3\end{aligned}$$

$$\begin{aligned}
\sigma^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \\
&= \frac{1}{10} \cdot ((2,5 - 4,3)^2 + (5,2 - 4,3)^2 + (4,5 - 4,3)^2 + (6,1 - 4,3)^2 + (1,6 - 4,3)^2 \\
&\quad + (3,6 - 4,3)^2 + (3,5 - 4,3)^2 + (7,0 - 4,3)^2 + (2,7 - 4,3)^2 + (6,3 - 4,3)^2) \\
&\approx 3,0 \\
\sigma &\approx 1,7
\end{aligned}$$

Die Zufallsvariable folgt also folgender Verteilung: $p(x|\omega_1) \sim N(4,3; 1,7)$

c) Bedingte Wahrscheinlichkeit nach Bayes:

$$P(\omega_c|x) = \frac{P(x|\omega_c)P(\omega_c)}{P(x)}$$

hierbei ist $P(x)$ gleich für $x \in \{1; 2\}$. Daher ist $\frac{P(x|\omega_c)P(\omega_c)}{p(x)}$ proportional zu $P(x|\omega_c)P(\omega_c)$.

Wir suchen also ein ω_c , das $P(x|\omega_c)P(\omega_c)$ maximiert.

Für eine neue Zufallszahl $x = 5.1$ und $c = 1$,

$$\begin{aligned}
P(\omega_1|x) &= \frac{P(x|\omega_1)P(\omega_1)}{P(x)} \\
P(x|\omega_1)P(\omega_1) &= \frac{1}{\sigma_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} P(\omega_1) \\
&= \frac{1}{1,7 \cdot \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{5,1-4,3}{1,7}\right)^2} \cdot 0,4 \\
&\approx 0,23 \cdot e^{-0,11} \cdot 0,4 \approx 0,1
\end{aligned}$$

für $c = 2$,

$$\begin{aligned}P(\omega_2|x) &= \frac{P(x|\omega_2)P(\omega_2)}{P(x)} \\P(x|\omega_2)P(\omega_2) &= \frac{1}{\sigma_2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2} P(\omega_2) \\&= \frac{1}{0,6 \cdot \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{5,1-5,1}{0,6}\right)^2} \cdot 0,6 \\&\approx 0,88 \cdot e^{0,0} \cdot 0,6 \approx 0,5\end{aligned}$$

Daraus folgt:

$$\begin{aligned}P(x|\omega_1)P(\omega_1) &\approx 0,1 \\P(x|\omega_2)P(\omega_2) &\approx 0,5\end{aligned}$$

Da $P(x|\omega_2)P(\omega_2)$ größer ist als $P(x|\omega_1)P(\omega_1)$ und $P(\omega_c|x) = \frac{P(x|\omega_c)P(\omega_c)}{P(x)}$ proportional zu $P(x|\omega_c)P(\omega_c)$, können wir auch schließen, dass $P(\omega_2|x) > P(\omega_1|x)$.

Deswegen wird $x = 5,1$ nach dem MAP-Kriterium die Klasse ω_2 zugewiesen.

Aufgabe 2: Fehlerwahrscheinlichkeit

Für zwei Klassen ω_1 und ω_2 mit den A-Priori-Wahrscheinlichkeiten $P(\omega_1) = \frac{1}{5}$ und $P(\omega_2) = \frac{4}{5}$ sind die folgenden Wahrscheinlichkeitsdichtefunktionen definiert:

$$P(x|\omega_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad \text{und} \quad P(x|\omega_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-2)^2}{2}}$$

- a) Wie hoch ist die Fehlerwahrscheinlichkeit in Abhängigkeit eines Schwellwerts θ , wenn sich eine Entscheidungsfunktion für ω_1 entscheidet wenn $x < \theta$ und ansonsten für ω_2 ?
- b) **Online-Frage Nr. 1:** Wie lautet der optimale Schwellwert θ_{opt} ?
- (i) 0,307
 - (ii) 0,653
 - (iii) 0,542
 - (iv) 0,451
 - (v) 1,000
- c) Berechnen Sie den kleinsten möglichen Klassifikationsfehler.
Hinweis: Formen Sie die Integrale aus der Funktion zur Berechnung der Fehlerwahrscheinlichkeit aus Teilaufgabe a) so um, dass sich deren Lösungen aus einer Tabelle zur Standardnormalverteilung $\Phi(z)$ ablesen lassen (z.B. unter http://de.wikipedia.org/wiki/Tabelle_Standardnormalverteilung).
- d) Sie haben eine Reihe von Messungen durchgeführt und die Messdaten mit einem Naiven Bayes-Klassifikator, bei dem Sie den Schwellwert auf das zuvor errechnete θ_{opt} gesetzt haben, klassifiziert. Sie stellen dabei fest, dass die tatsächliche Fehlerrate höher liegt als die in Teilaufgabe c) errechnete Fehlerwahrscheinlichkeit. Warum ist dies so?

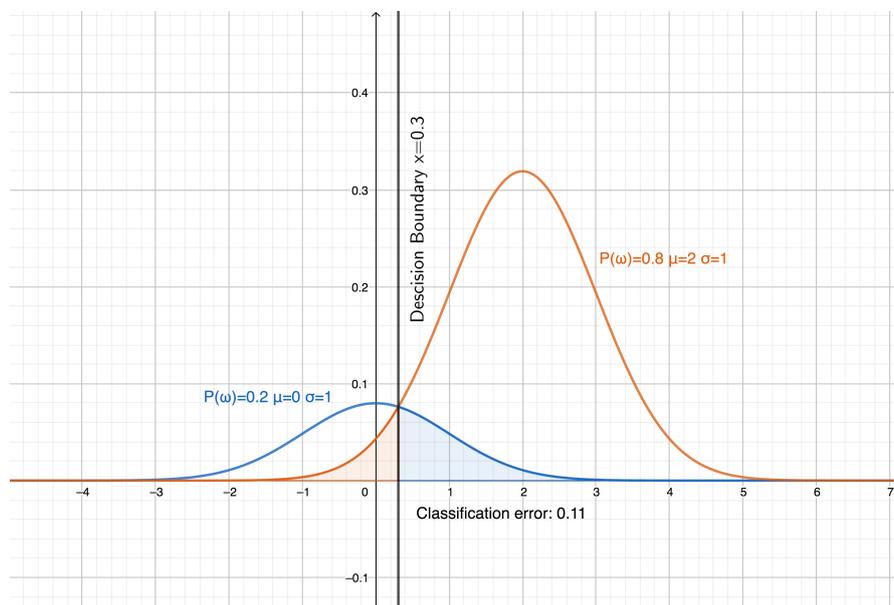
Lösung:

a)

$$\begin{aligned} P_{Fehler}(\theta) &= \int_{\theta}^{\infty} P(x|\omega_1)P(\omega_1)dx + \int_{-\infty}^{\theta} P(x|\omega_2)P(\omega_2)dx \\ &= \frac{1}{5} \int_{\theta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx + \frac{4}{5} \int_{-\infty}^{\theta} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-2)^2}{2}} dx \end{aligned}$$

- b) Der optimale Schwellwert θ_{opt} liegt an der Stelle an der sich die Fehlerwahrscheinlichkeitsfunktionen kreuzen. Der richtige Wert kann auch über eine grafische Darstellung am Schnittpunkt der beiden Funktionen abgelesen werden.

$$\begin{aligned}
P(\omega_1)P(x|\omega_1) &\stackrel{!}{=} P(\omega_2)P(x|\omega_2) \\
\frac{1}{5} \frac{1}{\sqrt{2\pi}} e^{-\frac{\theta_{opt}^2}{2}} &= \frac{4}{5} \frac{1}{\sqrt{2\pi}} e^{-\frac{(\theta_{opt}-2)^2}{2}} \\
1 e^{-\frac{\theta_{opt}^2}{2}} &= 4 e^{-\frac{(\theta_{opt}-2)^2}{2}} \\
\frac{1}{4} &= e^{\frac{\theta_{opt}^2 - (\theta_{opt}-2)^2}{2}} \\
0,25 &= e^{2\theta_{opt} - 2} \\
\ln 0,25 &= 2\theta_{opt} - 2 \\
\theta_{opt} &= \frac{2 + \ln 0,25}{2} \approx 0,307
\end{aligned}$$



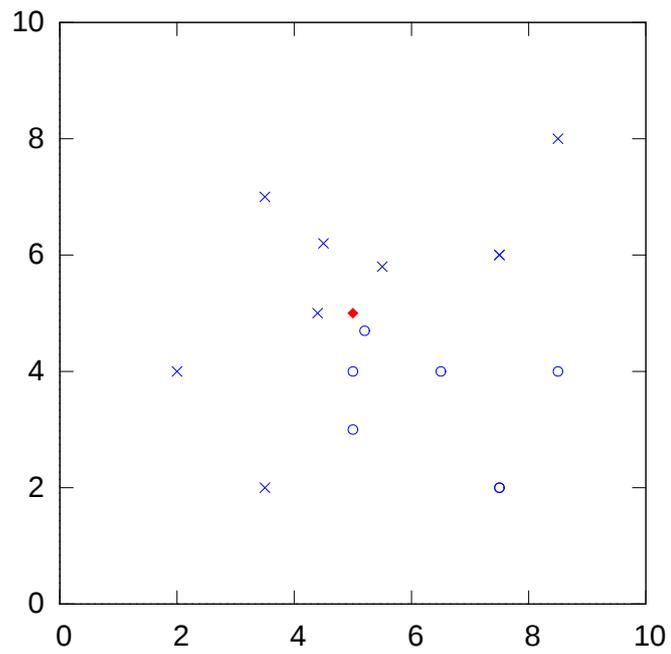
- c) Der kleinstmögliche Klassifikationsfehler wird erreicht, wenn der Schwellwert θ den optimalen Wert hat. Wir berechnen also $P_{fehler}(\theta_{opt})$. Die Standardverteilung ist: $\Phi_{0;1}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{1}{2}t^2} dt$.

$$\begin{aligned}
P_{fehler}(\theta) &= \int_{\theta}^{\infty} P(x|\omega_1)P(\omega_1)dx + \int_{-\infty}^{\theta} P(x|\omega_2)P(\omega_2)dx \\
&= \frac{1}{5} \int_{\theta}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx + \frac{4}{5} \int_{-\infty}^{\theta} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-2)^2}{2}} dx \\
&= \frac{1}{5} \int_{0,3}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx + \frac{4}{5} \int_{-\infty}^{0,3} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-2)^2}{2}} dx \\
&= \frac{1}{5} \cdot \frac{1}{\sqrt{2\pi}} \int_{0,3}^{\infty} e^{-\frac{x^2}{2}} dx + \frac{4}{5} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0,3-2} e^{-\frac{x^2}{2}} dx \\
&= \frac{1}{5} \cdot \left(1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0,3} e^{-\frac{x^2}{2}} dx\right) + \frac{4}{5} \cdot \left(1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1,7} e^{-\frac{x^2}{2}} dx\right) \\
&= \frac{1}{5} \cdot (1 - \Phi(0,3)) + \frac{4}{5} \cdot (1 - \Phi(1,7)) \\
&\approx \frac{1}{5} \cdot (1 - 0,62) + \frac{4}{5} \cdot (1 - 0,96) \\
&= \frac{1}{5} \cdot 0,38 + \frac{4}{5} \cdot 0,04 \\
&\approx 0,11
\end{aligned}$$

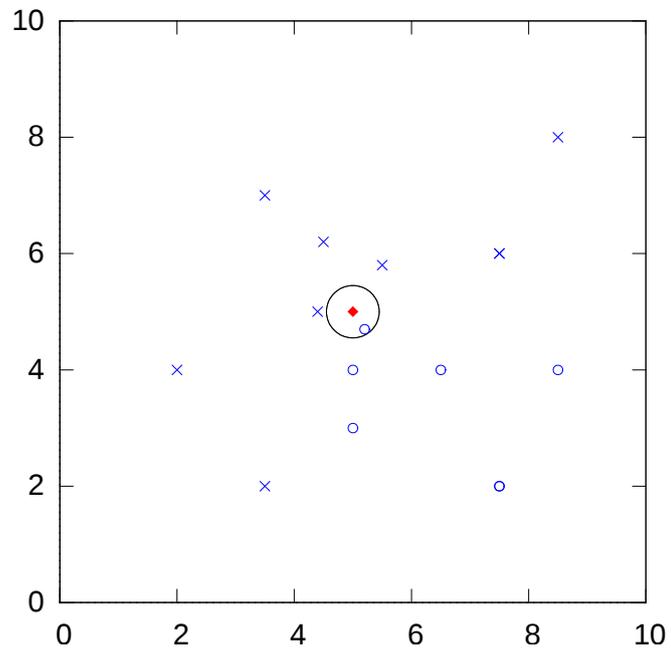
- d)
- Die Wahrscheinlichkeitsdichtefunktion ist immer nur eine Schätzung, die der realen Verteilung möglichst nahe kommen soll, sie aber nicht perfekt abbildet.
 - Zu wenige oder verzerrte Messdaten.

Aufgabe 3: *k*-Nearest-Neighbors

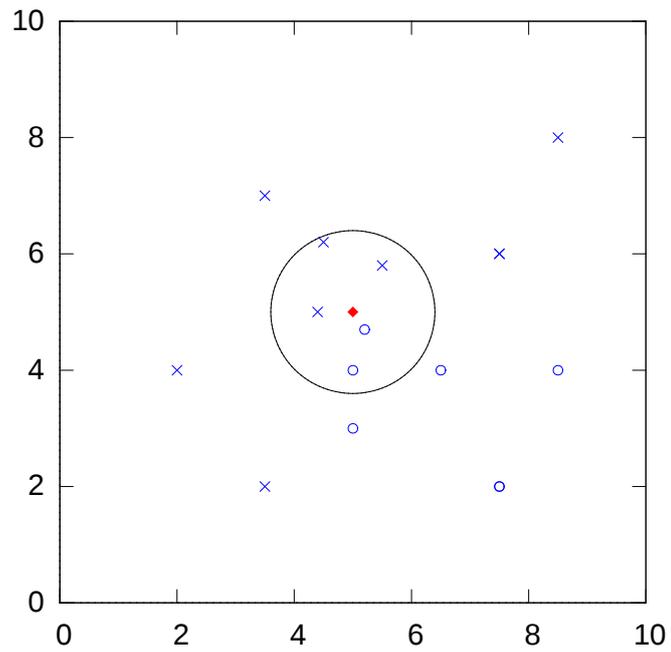
Gegeben sei der folgende Datensatz mit zwei Klassen (Kreise und Kreuze). Klassifizieren Sie das mit einer Raute markierte Merkmal an den Koordinaten (5, 5) mit dem *k*-Nearest-Neighbors Algorithmus für $k = 1$ und $k = 5$. Benutzen Sie dabei zur Gewichtung die euklidische Distanz.



Lösung für $k = 1$: Die Raute wird als Kreis klassifiziert.



Lösung für $k = 5$: Die Raute wird als Kreuz klassifiziert.

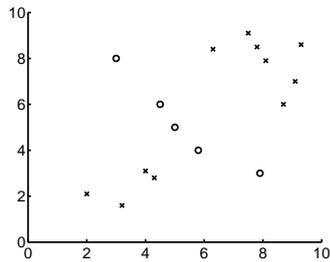


Online-Frage Nr. 2: Wie wird die Raute für $k = 5$ klassifiziert?

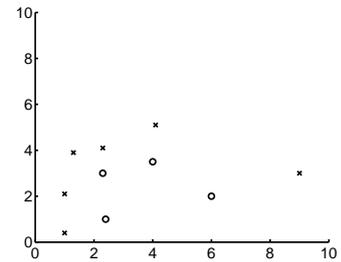
- (i) keine Klasse
- (ii) Kreis
- (iii) Kreuz

Aufgabe 4: Perzeptronen

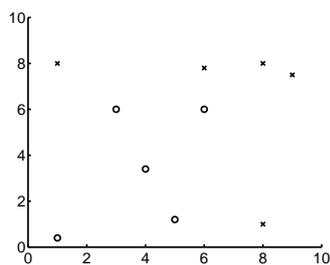
In den Grafiken (a) bis (d) sind jeweils Datenpunkte zweier Klassen eingezeichnet.



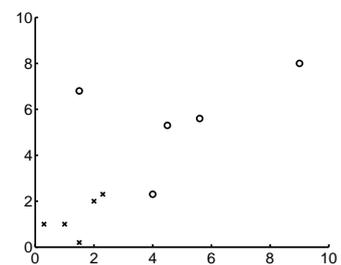
(a)



b)



(c)



d)

Online-Frage Nr. 3: Welche der folgenden Aussagen sind für die jeweils angegebene Abbildung wahr? Bitte alle richtigen Antworten im Online-Formular ankreuzen.

- (i) Abbildung a) Die beiden Klassen sind linear separierbar.
- (ii) Abbildung b) Die beiden Klassen lassen sich mit einem *Multi Layer Perceptron* mit einer versteckten Schicht, die aus zwei Neuronen besteht, trennen.
- (iii) Abbildung c) Nach einer Umrechnung in Polarkoordinaten (ohne Verschiebung) sind die beiden Klassen linear separierbar.
- (iv) Abbildung d) Die beiden Klassen lassen sich mit einem einfachen Perzeptron trennen.

Lösung:

- (i) Falsch.
- (ii) Richtig.
- (iii) Falsch.
- (iv) Richtig.

Aufgabe 5: Neuronale Netze

Gegeben seien die folgenden Punktwolken mit zweidimensionalen Merkmalsvektoren, die der positiven Klasse A und negativen Klasse B zugeordnet sind:

$$A = \{\vec{a}_1, \vec{a}_2\} = \left\{ \begin{pmatrix} -0.4 \\ 0.8 \end{pmatrix}; \begin{pmatrix} -0.8 \\ -0.2 \end{pmatrix} \right\} \text{ und } B = \{\vec{b}_1, \vec{b}_2\} = \left\{ \begin{pmatrix} 0.3 \\ 1.6 \end{pmatrix}; \begin{pmatrix} -0.4 \\ 1.5 \end{pmatrix} \right\}.$$

a) Berechnen Sie ein einfaches Perzeptron mit binärer Schwellwertfunktion, das die Klassen A und B trennt. Verwenden Sie den folgenden iterativen Algorithmus:

$$\begin{pmatrix} \vec{w}_{i+1} \\ b_{i+1} \end{pmatrix} = \begin{pmatrix} \vec{w}_i \\ b_i \end{pmatrix} - \sum_{\vec{e} \in E} \left[\delta_{\vec{e}} \cdot \begin{pmatrix} \vec{e} \\ 1 \end{pmatrix} \right]$$

- b_i entspricht dem Schwellenwert in Iterationsschritt i .
 - $\vec{w}_i = \begin{pmatrix} w_{i1} \\ w_{i2} \end{pmatrix}$ ist der Gewichtsvektor im Iterationsschritt i .
 - $E_i = \{\vec{e}_{i1}, \vec{e}_{i2}, \dots\} \subseteq \{\vec{a}_1, \vec{a}_2, \vec{b}_1, \vec{b}_2\}$ ist die Menge aller fehlerhaft klassifizierten Merkmalsvektoren in Iteration i .
 - $\delta_{\vec{e}}$ ist das Fehlersignal: $\delta_{\vec{e}} = -1$ falls $\vec{e} \in a$ und $\delta_{\vec{e}} = 1$ falls $\vec{e} \in b$.
 - Beginnen Sie die Iteration mit den Werten $\vec{w}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ und $b_0 = 0$ und skizzieren Sie die Trenngerade nach jeder Iteration i . Der Algorithmus terminiert in Iteration n , sobald alle Punkte richtig klassifiziert werden, d.h. $E_n = \emptyset$.
- b) **Onlinefrage Nr. 4:** Welchen Wert erhält man für $\{w_{n1}, w_{n2}, b_n\}$ wenn der Algorithmus terminiert?
- i) 3.2, -0.4, -1
 - ii) -0.5, -0.9, 1
 - iii) -0.9, -1.9, 2
 - iv) -1.0, -2.0, 2
 - v) 0.5, 2.8, 0

Lösung:

a) Die lineare Diskriminierungsfunktion eines einfachen Perzeptrons mit zwei Neuronen hat allgemein die Form: $g(x) = w_1x_1 + w_2x_2 + b$.

Die Klassenzuordnung erfolgt über den Wert von $g(x)$: Falls $g(x) > 0 \implies$ Klasse A und falls $g(x) < 0 \implies$ Klasse B. Die Trenngerade zwischen den beiden Klassen verläuft entsprechend bei $g(x) = 0$.

Zur Zeichnung in ein Achsenkreuz mit den beiden Achsen x_1 und x_2 wird die Trenngerade bei $g(x) = 0$ nach x_2 aufgelöst:

$$\begin{aligned}w_1x_1 + w_2x_2 + b &= 0 \\w_2x_2 &= -w_1x_1 - b \\x_2 &= -\frac{w_1}{w_2}x_1 - \frac{1}{w_2}b\end{aligned}$$

Als Trenngerade für die initialen Werte ergibt sich aus $g_0(x) = x_1 - x_2$ demnach die Geradengleichung $x_2 = x_1$ für ein zweidimensionales Achsenkreuz.

Der Algorithmus durchläuft die folgenden Iterationsschritte:

$$1. \vec{w}_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}; b_0 = 0 \implies g_0(x) = x_1 - x_2$$

$$g_0(\vec{a}_1) = -1.2 \stackrel{!}{\implies} B \text{ (Fehlklassifikation)}$$

$$g_0(\vec{a}_2) = -0.6 \stackrel{!}{\implies} B \text{ (Fehlklassifikation)}$$

$$g_0(\vec{b}_1) = -1.3 \implies B$$

$$g_0(\vec{b}_2) = -1.9 \implies B$$

Damit sind zwei Punkte aus Klasse A falsch klassifiziert, $E_0 = \{\vec{a}_1, \vec{a}_2\}$.

Daraus ergibt sich nach dem ersten Iterationsschritt:

$$\begin{aligned}\begin{pmatrix} \vec{w}_1 \\ b_1 \end{pmatrix} &= \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} - (-1)\vec{a}_1 - (-1)\vec{a}_2 \\ &= \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} - (-1) \begin{pmatrix} -0.4 \\ 0.8 \\ 1 \end{pmatrix} - (-1) \begin{pmatrix} -0.8 \\ -0.2 \\ 1 \end{pmatrix} = \begin{pmatrix} -0.2 \\ -0.4 \\ 2 \end{pmatrix} \\ \implies g_1(x) &= -0.2x_1 - 0.4x_2 + 2\end{aligned}$$

Geradengleichung für die Zeichnung: $x_2 = -0.5x_1 + 5$

$$2. g_1(\vec{a}_1) = 1.76 \implies A$$

$$g_1(\vec{a}_2) = 2.24 \implies A$$

$$g_1(\vec{b}_1) = 1.3 \stackrel{!}{\implies} A \text{ (Fehlklassifikation)}$$

$$g_1(\vec{b}_2) = 1.48 \stackrel{!}{\implies} A \text{ (Fehlklassifikation)}$$

Damit sind zwei Punkte aus Klasse B falsch klassifiziert: $E_1 = \{\vec{b}_1, \vec{b}_2\}$

$$\begin{aligned}\begin{pmatrix} \vec{w}_2 \\ b_2 \end{pmatrix} &= \begin{pmatrix} -0.2 \\ -0.4 \\ 2 \end{pmatrix} - (1)\vec{b}_1 - (1)\vec{b}_2 = \begin{pmatrix} -0.2 \\ -0.4 \\ 2 \end{pmatrix} - (1) \begin{pmatrix} 0.4 \\ 1.6 \\ 1 \end{pmatrix} - (1) \begin{pmatrix} -0.4 \\ 1.5 \\ 1 \end{pmatrix} = \begin{pmatrix} -0.1 \\ -3.5 \\ 0 \end{pmatrix} \\ \implies g_2(x) &= -0.1x_1 - 3.5x_2\end{aligned}$$

Geradengleichung für die Zeichnung: $x_2 = \frac{-x_1}{35}$

3. $g_2(\vec{a}_1) = -2.76 \stackrel{!}{\implies} A$
 $g_2(\vec{a}_2) = 0.78 \implies A$
 $g_2(\vec{b}_1) = -5.63 \implies B$
 $g_2(\vec{b}_2) = -5.21 \implies B$

$$\begin{pmatrix} \vec{w}_3 \\ b_3 \end{pmatrix} = \begin{pmatrix} -0.5 \\ -2.7 \\ 1 \end{pmatrix}$$

Geradengleichung für die Zeichnung: $x_2 = \frac{-5x_1+10}{27}$

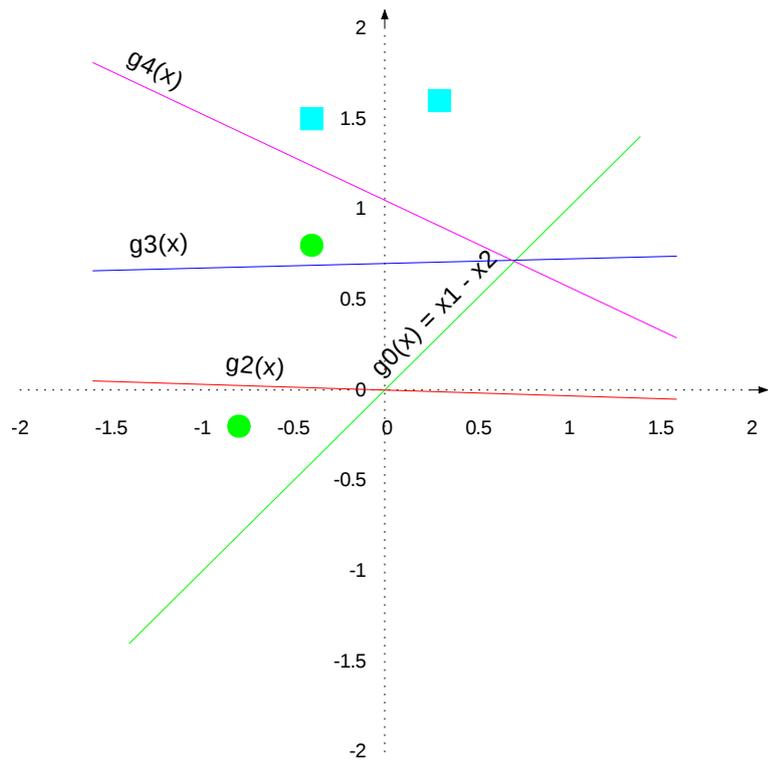
4. $g_3(\vec{a}_1) = -0.96 \stackrel{!}{\implies} A$
 $g_3(\vec{a}_2) = 1.94 \implies A$
 $g_3(\vec{b}_1) = -4.47 \implies B$
 $g_3(\vec{b}_2) = -2.85 \implies B$

$$\begin{pmatrix} \vec{w}_4 \\ b_4 \end{pmatrix} = \begin{pmatrix} -0.9 \\ -1.9 \\ 2 \end{pmatrix}$$

Geradengleichung für die Zeichnung: $x_2 = \frac{-9x_1+20}{19}$

5. $g_4(\vec{a}_1) = 0.84 \implies A$
 $g_4(\vec{a}_2) = 3.1 \implies A$
 $g_4(\vec{b}_1) = -1.31 \implies B$
 $g_4(\vec{b}_2) = -0.49 \implies B$

Alle Trainingsdatenpunkte werden damit korrekt klassifiziert: $E_2 = \emptyset \implies$ der Algorithmus terminiert mit den Werten $(\vec{w}_4, b_4) = (-0.9, -1.9, 2.0)^T$.



Lösungsblatt 3

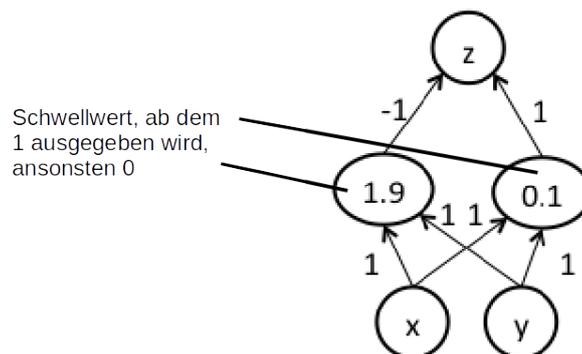
Maschinelles Lernen und Spracherkennung

Aufgabe 1: Neuronales Netz: XOR-Problem

- a) Zeichnen Sie ein neuronales Netz mit genau einer versteckten Schicht, welches das XOR-Problem löst. Nehmen Sie hierfür 2 binäre Eingänge und einen binären Ausgang an, die jeweils einen der Werte 0 oder 1 annehmen können. Die Aktivierungsfunktion ist die Sigmoidfunktion.

Lösung:

Mögliche Lösung:



Onlinefrage Nr. 1: Wieviele versteckte Neuronen muss das Netz aus Aufgabe 1 a) mindestens enthalten (d.h., wie viele Knoten enthält die versteckte Schicht), um das XOR-Problem zu lösen? (Hinweis: direkte Verbindungen zwischen Eingabe und Ausgabe seien explizit *nicht* erlaubt).

- i) 1
- ii) 2
- iii) 3
- iv) 4
- v) 5

Lösung:

- ii) 2 (für den Fall das direkte Verbindungen nicht erlaubt sind, ansonsten lautet die Lösung 1).

Aufgabe 2: Neuronales Netz: Autoencoder

Ein Autoencoder ist ein künstliches neuronales Netzwerk, welches eine Repräsentation (ein Encoding) eines Datensatzes lernt, typischerweise mit dem Ziel der Dimensionen-Reduktion. Meist besteht ein Autoencoder aus einem Multilayer-Perzeptron (MLP) mit gleicher Zahl von Eingabe- und Ausgabeknoten, und einer versteckten Schicht von geringerer Dimension. Dieses Netz wird dann so trainiert, dass es die Eingabe wieder möglichst genau rekonstruiert. Mit anderen Worten, im Training ist der erwünschte Ausgabe-Vektor gleich dem Eingabevektor.

Welche Eigenschaften gelten für einen Autoencoder?

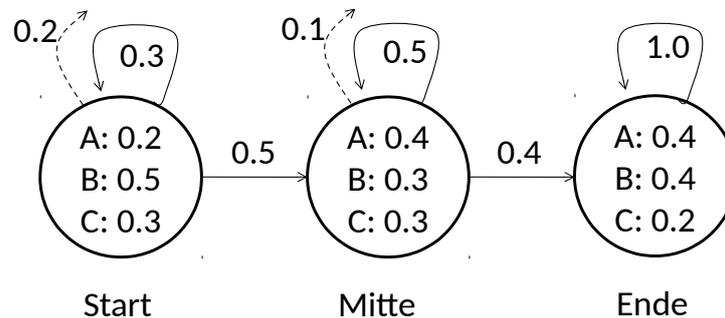
- i) unüberwacht
- ii) überwacht
- iii) nicht-parametrisch
- iv) parametrisch

Lösung:

i) and iii)

Aufgabe 3: HMM - Forward- / Viterbi-Algorithmus

Gegeben sei das Modell λ wie folgt:



Hinweis: Die Zustandsübergänge summieren sich nicht zu 1, dies ist also kein vollständig beschriebenes HMM. Die implizite Annahme ist hier, dass es noch andere, nicht gezeigte Zustände und entsprechende Übergänge gibt, die jedoch die beobachtete Sequenz nicht erzeugen können und somit für die Aufgabe keine Rolle spielen.

- a) Berechnen Sie die Wahrscheinlichkeit der Zeichenketten $O_1 = AAB$ und $O_2 = BAB$ mit dem Forward-Algorithmus, wobei im Start-Zustand begonnen, im Ende-Zustand terminiert und im Start-Zustand q_0 noch kein Symbol emittiert wird (d.h., es kommt immer zuerst ein Zustandsübergang, und erst danach eine Beobachtung).

b) Welche Zustandskette $Q = (q_0 q_1 q_2 q_3)$ liefert die maximale Wahrscheinlichkeit $P(O, Q|\lambda)$ für die Ausgabe O_2 mit $q_0 = Start$ wobei im Startzustand q_0 noch kein Symbol emittiert wird? (Hinweis: Annahme ist wieder, das im Startzustand begonnen und im Endzustand terminiert wird).

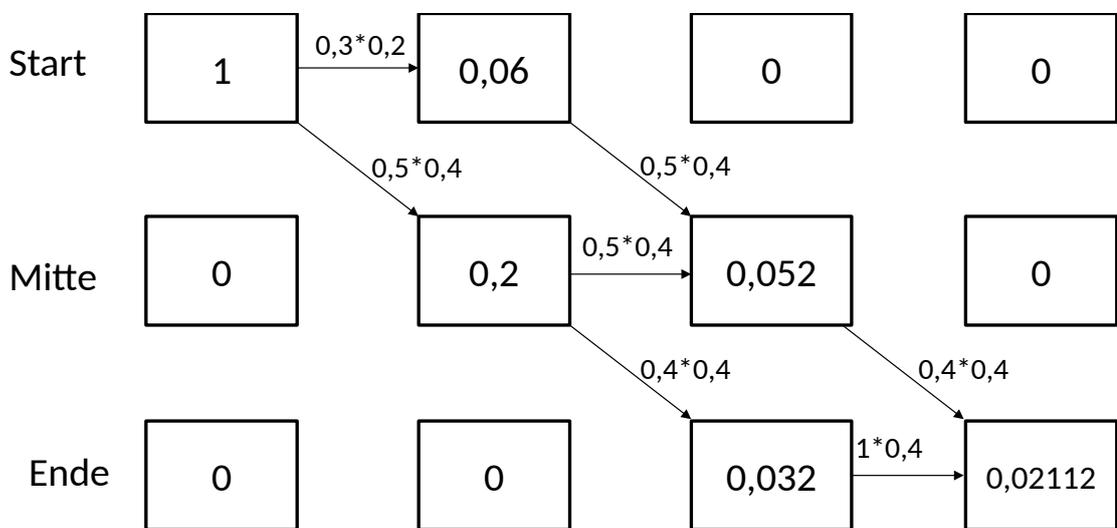
Hinweis: Wenden Sie hier den Viterbi-Algorithmus an.

c) **Onlinefrage Nr. 2:** Welche Zeichenkette aus Teilaufgabe a) ist wahrscheinlicher?

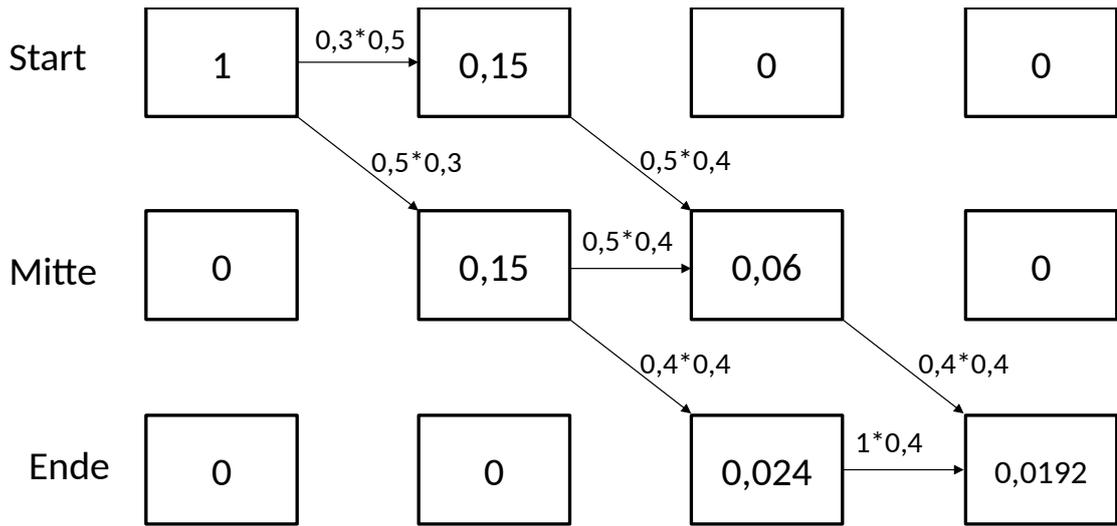
- i) O_1
- ii) O_2
- iii) O_1 und O_2

Lösung:

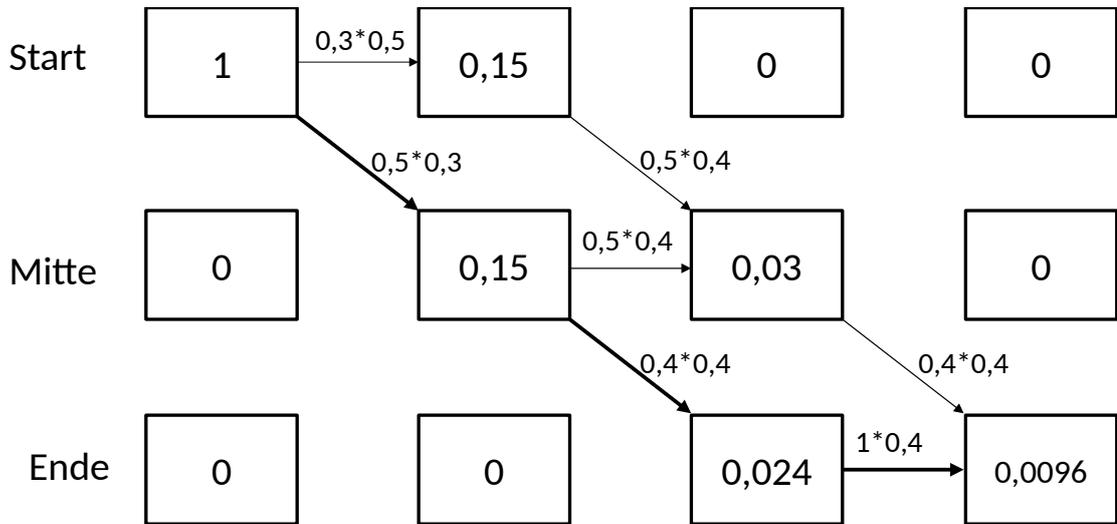
a) $P(O_1, Q|\lambda) = 0.02112$



$P(O_2, Q|\lambda) = 0.0192$



b) Der wahrscheinlichste Pfad für die Beobachtungssequenz O_2 führt entlang der fett gedruckten Pfeile.



c) Die Folge O_1 ist wahrscheinlicher, da $P(O_1, Q|\lambda) = 0.02112 > P(O_2, Q|\lambda) = 0.0192$

Aufgabe 4: Sprachmodelle

Ein 3-gram Sprachmodell einer formalen Sprache mit dem Vokabular

$V = \{Karlsruher, Institut, für, Technologie\}$ enthalte die folgenden Wahrscheinlichkeiten.

w_{i-2}	w_{i-1}	$w_i = \text{Karlsruher}$	$w_i = \text{Institut}$	$w_i = \text{für}$	$w_i = \text{Technologie}$	$w_i = \text{</S>}$
n/a	<S>	0,5	0,25	0,1	0,1	0,05
<S>	Karlsruher	0	0,8	0,05	0,05	0,1
<S>	Institut	1	0	0	0	0
<S>	für	0	0,9	0	0	0,1
<S>	Technologie	0	0,7	0,2	0	0,1
Karlsruher	Karlsruher	0,2	0,2	0,2	0,2	0,2
Karlsruher	Institut	0	0	0,5	0,5	0
Karlsruher	für	0	0,6	0,2	0,1	0,1
Karlsruher	Technologie	0	0,4	0,4	0,1	0,1
Institut	Karlsruher	0	0	0,4	0,4	0,2
Institut	Institut	0,2	0,2	0,2	0,2	0,2
Institut	für	0,1	0,1	0,1	0,1	0,6
Institut	Technologie	0,05	0,05	0,6	0,1	0,2
für	Karlsruher	0,1	0,6	0,1	0,1	0,1
für	Institut	0,3	0	0,1	0,2	0,4
für	für	0	0	0	0	0
für	Technologie	0,2	0,2	0,2	0,2	0,2
Technologie	Karlsruher	0	0,4	0,1	0,1	0,4
Technologie	Institut	0,4	0	0	0	0,6
Technologie	für	0,1	0,4	0	0	0,5
Technologie	Technologie	0,2	0,2	0,2	0,2	0,2

Hierbei steht das Symbol <S> für den Anfang des Satzes und </S> für das Satzende. Berechnen Sie die Wahrscheinlichkeiten der folgenden Sätze und die dazugehörigen Perplexitäten des Sprachmodells.

Hinweis:

Als „Anzahl der Wörter“ bei der Berechnung der Perplexität in der Spracherkennung wird oft die Anzahl der Wortübergänge benutzt (bei n-Grammen mit $n > 1$), da diese letztendlich entscheidend für die Wahrscheinlichkeit der Wortsequenz sind. Benutzen Sie zur Berechnung der Perplexität die Formeln für Perplexität und Entropie aus diesem Hinweis und verwenden Sie die Anzahl der Wortübergänge als n in der Berechnung der Entropie. <S> und </S> werden dabei als Wörter angesehen.

Formel für die Berechnung der Wahrscheinlichkeit mit 3-Grammen:

$$\hat{P}(W) = P(w_1, w_2, \dots, w_n) = \prod_{i=1}^n P(w_i | w_{i-2} w_{i-1})$$

Formel für die Berechnung der Perplexität:

$$PP(W) = 2^{\hat{H}(W)} \text{ mit } \hat{H}(W) = -\frac{1}{n} \sum_{i=1}^n \log_2 P(w_i | w_{i-2} w_{i-1})$$

a) „Karlsruher Institut für Technologie“

Lösung:

Formel für die Berechnung der Wahrscheinlichkeit mit 3-Grammen:

$$\hat{P}(W) = P(w_1, w_2, \dots, w_n) = \prod_{i=1}^n P(w_i | w_{i-2} w_{i-1})$$

$$\begin{aligned}
& P(\langle s \rangle \text{ Karlsruher Institut für Technologie } \langle /s \rangle) \\
&= P(\text{Karlsruher} \mid n/a \langle s \rangle) * P(\text{Institut} \mid \langle s \rangle \text{ Karlsruher}) * P(\text{für} \mid \text{Karlsruher Institut}) \\
&* P(\text{Technologie} \mid \text{Institut für}) * P(\langle /s \rangle \mid \text{für Technologie}) \\
&= 0.5 * 0.8 * 0.5 * 0.1 * 0.2 \\
&= 0.004
\end{aligned}$$

Formel für die Berechnung der Perplexität:

$$PP(W) = 2^{\hat{H}(W)} \text{ mit } \hat{H}(W) = -\frac{1}{n} \sum_{i=1}^n \log_2 P(w_i \mid w_{i-2} w_{i-1})$$

Umformung:

$$\begin{aligned}
PP(W) &= 2^{-\frac{1}{n} \sum_{i=1}^n \log_2 P(w_i \mid w_{i-2} w_{i-1})} = 2^{\sum_{i=1}^n -\frac{1}{n} \log_2 P(w_i \mid w_{i-2} w_{i-1})} = 2^{\sum_{i=1}^n \log_2 P(w_i \mid w_{i-2} w_{i-1})^{-\frac{1}{n}}} = \\
&= 2^{\log_2 \prod_{i=1}^n P(w_i \mid w_{i-2} w_{i-1})^{-\frac{1}{n}}} = \prod_{i=1}^n P(w_i \mid w_{i-2} w_{i-1})^{-\frac{1}{n}}
\end{aligned}$$

$$\begin{aligned}
PP(\langle S \rangle \text{ Karlsruher Institut für Technologie } \langle /S \rangle) &= \\
P(\langle S \rangle \text{ Karlsruher Institut für Technologie } \langle /S \rangle)^{-1/5} &= 0,004^{-1/5} \approx 3,0171
\end{aligned}$$

b) „Karlsruher für Technologie Institut“

Lösung:

$$\begin{aligned}
& P(\langle s \rangle \text{ Karlsruher für Technologie Institut } \langle /s \rangle) \\
&= P(\text{Karlsruher} \mid n/a \langle s \rangle) * P(\text{für} \mid \langle s \rangle \text{ Karlsruher}) * P(\text{Technologie} \mid \text{Karlsruher für}) \\
&* P(\text{Institut} \mid \text{für Technologie}) * P(\langle /s \rangle \mid \text{Technologie Institut}) \\
&= 0.5 * 0.05 * 0.1 * 0.2 * 0.6 \\
&= 0.0003
\end{aligned}$$

$$\begin{aligned}
PP(\langle S \rangle \text{ Karlsruher für Technologie Institut } \langle /S \rangle) &= \\
P(\langle S \rangle \text{ Karlsruher für Technologie Institut } \langle /S \rangle)^{-1/4} &= 0,0003^{-1/5} \approx 5.065
\end{aligned}$$

c) „Institut Karlsruher Technologie für“

Lösung:

$$\begin{aligned}
& P(\langle s \rangle \text{ Institut Karlsruher Technologie für } \langle /s \rangle) \\
&= P(\text{Institut} \mid n/a \langle s \rangle) * P(\text{Karlsruher} \mid \langle s \rangle \text{ Institut}) * P(\text{Technologie} \mid \text{Institut Karlsruher}) \\
&* P(\text{für} \mid \text{Karlsruher Technologie}) * P(\langle /s \rangle \mid \text{Technologie für}) \\
&= 0.25 * 1 * 0.4 * 0.4 * 0.5 \\
&= 0.02
\end{aligned}$$

$$\begin{aligned}
PP(\langle S \rangle \text{ Institut Karlsruher Technologie für } \langle /S \rangle) &= \\
P(\langle S \rangle \text{ Institut Karlsruher Technologie für } \langle /S \rangle)^{-1/5} &= 0,02^{-1/5} \approx 2,1867
\end{aligned}$$

d) „Für Institut Institut Technologie“

Lösung:

$$\begin{aligned}
& P(\langle s \rangle \text{ für Institut Institut Technologie } \langle /s \rangle) \\
&= P(\text{für} \mid n/a \langle s \rangle) * P(\text{Institut} \mid \langle s \rangle \text{ für}) * P(\text{Institut} \mid \text{für Institut}) \\
&* P(\text{Technologie} \mid \text{Institut Institut}) * P(\langle /s \rangle \mid \text{Institut Technologie}) \\
&= 0.1 * 0.9 * 0.2 * 0.2
\end{aligned}$$

= 0

$PP(\langle S \rangle \text{ für Institut Institut Technologie } \langle /S \rangle) =$
 $P(\langle S \rangle \text{ für Institut Institut Technologie } \langle /S \rangle)^{-1/5} \rightarrow ERROR$

Anmerkung: Die Perplexität kann hier nicht berechnet werden, da der Satz nicht im Sprachmodell erlaubt ist (z.B. eines der 3-Gramme wurde im Training nie gesehen). In der Praxis bekommen daher nicht gesehene 3-Gramme eine kleine „Backoff“ Wahrscheinlichkeit zugewiesen, entweder fix oder durch Interpolation mit Bi-Grammen, wenn diese nicht gesehen wurden mit Uni-Grammen.

Onlinefrage Nr. 3: Geben Sie die Sätze a)-d) nach absteigender Wahrscheinlichkeit sortiert an.

- i) *abcd*
- ii) *cabd*
- iii) *badc*
- iv) *cdba*
- v) *cdab*

Lösung:

Aus den Aufgaben 3a)-d) wissen wir: $P(a) = 0,004$; $P(b) = 0,0003$; $P(c) = 0,02$; $P(d) = 0 \rightarrow$
cabd

Aufgabe 5: Wortfehlerrate

Zur Bewertung eines Spracherkenners wird oft die Wortfehlerrate (Word Error Rate, WER) herangezogen:

$$WER = \frac{N_{sub} + N_{del} + N_{ins}}{N}$$

Dabei ist N die Anzahl der Wörter in der Referenz und N_{op} sind die Anzahl der Wörter, die in der Referenz mit Operation *op* bearbeitet werden müssen, um die Hypothese zu erhalten.

- Auslassung / Deletion (del): Ein Wort wird aus der Referenz entfernt.
- Einfügungen / Insertion (ins): Ein Wort wird in die Referenz eingefügt.
- Verwechslung / Substitution (sub): Ein Wort in der Referenz wird gegen ein Wort aus der Hypothese ausgetauscht.

Berechnen Sie für die folgenden Sätze die Wortfehlerrate, wenn für die automatisch erkannte Hypothese HYP die Referenz REF gestellt wird. Bei der Rechnung sollen Satzzeichen vernachlässigt werden.

- a) HYP „bluten Morgen wie geht es hier“
REF „guten Morgen wie geht es“

Lösung:

sub, match, match, match, ins
 $\Rightarrow WER \text{ is } \frac{2}{5} = 40\%$

- b) HYP „es das Maus ist und groß“
REF „das Haus ist blau und groß“

Lösung:

ins, match, sub, match, del, match, match
⇒ WER is $\frac{3}{6} = 50\%$

- c) HYP „die automatische Sprache Erkennung ist schwierig es“
REF „automatische Spracherkennung ist schwierig“

Lösung:

ins, match, sub, ins, match, match, ins
⇒ WER is $\frac{4}{4} = 100\%$

Onlinefrage Nr. 4: Berechnen Sie die Wortfehlerrate für die folgende Hypothese und Referenz. Auch hier sollen die Satzzeichen für die Berechnung vernachlässigt werden.

- d) HYP „Wen essen Sommer wie der Wahn ist, gehen vier Eisen.“
REF „Wenn es im Sommer wieder warm ist, gehen wir Eis essen.“

- i) 11/9
- ii) 9/10
- iii) 9/11
- iv) 10/11

Lösung:

sub, sub, del, match, sub, ins, sub, match, match, sub, sub, del
⇒ WER is $\frac{9}{11}$

Aufgabe 6: Programmieraufgabe: Dynamische Programmierung

Gegeben sind die korrekte Sequenz (Referenz) „, wenn es im Juni viel donnert kommt ein trüber Sommer“ und fünf Hypothesen:

- 1) im Juni viel Sonne kommt einen trüberen Sommer
 - 2) viel Donner im Juni einen trüben Sommer bringt
 - 3) Juni Donner einen Sommer
 - 4) im Juni viel Donner bringt einen trüben Sommer
 - 5) wenns im Juno viel Donner gibts einen trüben Sommer
- a) Schreiben Sie ein Programm, um die gegebenen Hypothesen mit der korrekten Sequenz auf Wortebene (Editierdistanz auf Wortebene, englisch *word edit distance*) zu vergleichen. Implementieren Sie dabei eine Zurückverfolgung des Pfades und vergeben Sie Strafpunkte, um dadurch die ähnlichste Hypothese auszuwählen. Dabei sollen übereinstimmende Wörter 0 Strafpunkte bekommen, und substitutions (*Ersetzungen*), insertions (*Einfügungen*), und deletions (*Auslassungen*) jeweils 1 Strafpunkt. Achten Sie darauf, dass ihr Algorithmus *case-sensitive* arbeitet, d.h. “Donner” und “donner” als 2 unterschiedliche Wörter betrachtet. Welche der 5 Hypothesen ist der Referenz am ähnlichsten?

Lösung:

Beispiel Python-Code

(Quelle: en.wikibooks.org/wiki/Algorithm_Implementation/Strings/Levenshtein_distance (erweitert um Pfadzurückverfolgung)):

```
def levenshtein(s1, s2, backtracking=False):
    if len(s1) < len(s2):
        return levenshtein(s2, s1, backtracking=backtracking)
    if len(s2) == 0:
        return len(s1)

    path = []

    previous_row = range(len(s2) + 1)
    path.append(['start'] + ['del'] * len(s2))

    for i, c1 in enumerate(s1):
        current_row = [i + 1]
        current_row_path = ['ins']
        for j, c2 in enumerate(s2):
            insertions = previous_row[j + 1] + 1
            deletions = current_row[j] + 1
            substitutions = previous_row[j] + (c1 != c2)
            current_row.append(min(insertions, deletions, substitutions))

            if insertions < deletions and insertions < substitutions:
                current_row_path.append('ins')
            elif deletions < insertions and deletions < substitutions:
                current_row_path.append('del')
```

```

    elif c1 == c2:
        current_row_path.append('match')
    else:
        current_row_path.append('sub')

    previous_row = current_row
    path.append(current_row_path)

    if backtracking:
        return previous_row[-1], path

    return previous_row[-1]

```

Distanzen auf Wortebene: 5, 8, 8, 6, 7. Somit ist Hypothese 1 der Referenz am ähnlichsten.

- b) Modifizieren Sie Ihr Programm, so dass die Ähnlichkeit auf Buchstabenebene bestimmt wird (*character edit distance*). Dazu soll jeder Buchstabe inklusive Leerzeichen als eigenes "Wort" behandelt werden. Welche der 5 Hypothesen ist nun der Referenz am ähnlichsten?

Lösung:

Anstatt Wörter in einer Liste übergeben, String übergeben (String ist eine Liste von Buchstaben). Distanzen auf Wortebene: 15, 33, 30, 18, 13. Somit ist Hypothese 5 der Referenz am ähnlichsten.

- c) **Onlinefrage Nr. 5:** Nun sei die Anzahl der Strafpunkte für substitutions (*Ersetzungen*) 2 (statt 1), die anderen Strafpunkte bleiben wie zuvor. Welche Hypothese ist nun der Referenz am ähnlichsten (gemessen in Editierdistanz auf Wortebene), und was ist die entsprechende Editierdistanz?
- i) Hypothese 4 (Editierdistanz 6)
 - ii) Hypothese 1 (Editierdistanz 5)
 - iii) Hypothese 2 (Editierdistanz 5)
 - iv) Hypothese 3 (Editierdistanz 8)
 - v) Hypothese 1 (Editierdistanz 8)
 - vi) Hypothese 2 (Editierdistanz 6)

Lösung: v) Hypothese 1 (Editierdistanz 8)

Hinweis:

Die Onlinefrage Nr. 5 kann auch durch manuelles Berechnen der Punktzahlen beantwortet werden, jedoch dürfte eine Implementierung schneller und sicherer sein, da man sich bei diesem Algorithmus leicht verrechnen kann.

Anmerkung zur Dynamischen Programmierung

Man kann die dynamischen Programmierung als eine Art Viterbi-Algorithmus ansehen, bei dem statt dem Pfad mit der maximalen Wahrscheinlichkeit der Pfad mit dem höchsten Score gesucht wird.

Die Dynamische Programmierung unterteilt sich in drei Schritte:

1. Die Generierung einer Matrix mit $M + 1$ Spalten auf $N + 1$ Zeilen, wobei M die Anzahl der Wörter der korrekten Sequenz ist, und N die Anzahl der Wörter in der Hypothese. Die erste Zeile

und Spalte können bereits mit der Bewertung für ausschließlich Auslassungen bzw. Einfügungen gefüllt werden, siehe Beispiel:

		im	Juni	viel	Donner	bringt	einen	trüben	Sommer
	0	1	2	3	4	5	6	7	8
im	1								
Juni	2								
viel	3								
Donner	4								
bringt	5								
einen	6								
trüben	7								
Sommer	8								

2. Das Füllen der Matrixfelder (engl. *scoring*) mit dem maximalen Score.
3. Das Zurückverfolgen des Pfades (engl. *alignment*).

Musterlösung zu Übungsblatt 4

Grundlagen der Bildverarbeitung

Aufgabe 1 – Farbmodell

1.a

$$\mathbf{r} = (1 \ 0 \ 0)^T: c = \arccos(1) = 0^\circ \Rightarrow H = 0^\circ, S = 1, I = 1/3$$

$$\mathbf{g} = (0 \ 1 \ 0)^T: c = \arccos(-0,5) = 120^\circ \Rightarrow H = 120^\circ, S = 1, I = 1/3$$

$$\mathbf{b} = (0 \ 0 \ 1)^T: c = \arccos(-0,5) = 120^\circ \Rightarrow H = 240^\circ, S = 1, I = 1/3$$

1.b

$$\mathbf{f} = (1 \ 1 \ 1)^T: H = \text{undefiniert}, S = 0, I = 1$$

Für weiß ($S = 0$) und schwarz ($I = 0$) ist der H-Wert undefiniert.

Onlinefrage Nr. 1:

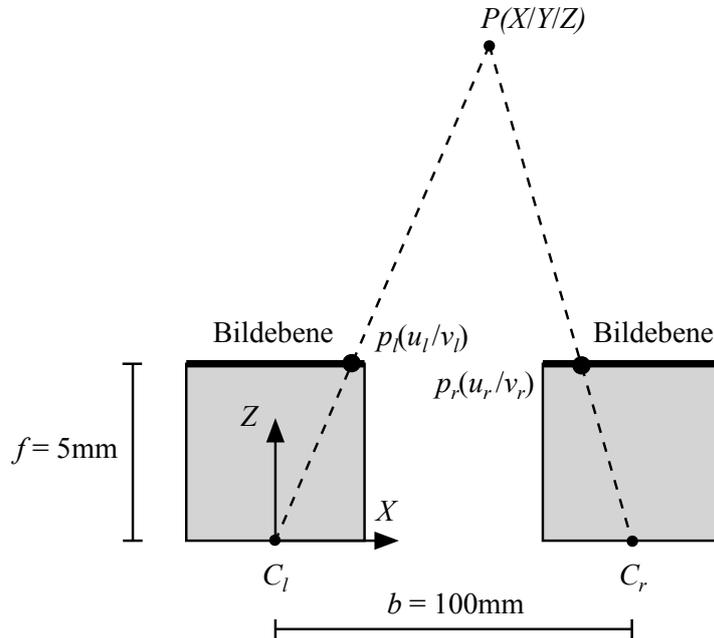
Den Unterschied zwischen den Farben Rot und Pastell-Rot kann man anhand des Wertes Saturation erkennen.

Aufgabe 2 – Lochkameramodell

2.a

Ein Pixel ist $\frac{6,4 \text{ mm}}{640} = \frac{4,8 \text{ mm}}{480} = 0,01\text{mm}$ hoch und breit \Rightarrow Umrechnungsfaktor ist $\frac{1}{0,01} \frac{\text{Pixel}}{\text{mm}} = 100 \frac{\text{Pixel}}{\text{mm}}$.

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{f}{Z} \begin{pmatrix} X \\ Y \end{pmatrix} = \frac{5 \text{ mm}}{1000 \text{ mm}} \begin{pmatrix} -300 \text{ mm} \\ 500 \text{ mm} \end{pmatrix} = \frac{500 \text{ Pixel}}{1000 \text{ mm}} \begin{pmatrix} -300 \text{ mm} \\ 500 \text{ mm} \end{pmatrix} = \begin{pmatrix} -150 \text{ Pixel} \\ 250 \text{ Pixel} \end{pmatrix}$$



4.2.b

$$g_l : \vec{x} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{Z}{f} \begin{pmatrix} u_l \\ v_l \\ f \end{pmatrix} = \frac{Z}{f} \begin{pmatrix} 0 \text{ Pixel} \\ -125 \text{ Pixel} \\ 5 \text{ mm} \end{pmatrix} = \frac{Z}{f} \begin{pmatrix} 0 \text{ mm} \\ -1,25 \text{ mm} \\ 5 \text{ mm} \end{pmatrix} = r \begin{pmatrix} 0 \text{ mm} \\ -1,25 \text{ mm} \\ 5 \text{ mm} \end{pmatrix}$$

$$g_r : \vec{x} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \frac{Z}{f} \begin{pmatrix} u_r \\ v_r \\ f \end{pmatrix} = \frac{Z}{f} \begin{pmatrix} -250 \text{ Pixel} \\ -125 \text{ Pixel} \\ 5 \text{ mm} \end{pmatrix} = \frac{Z}{f} \begin{pmatrix} -2,5 \text{ mm} \\ -1,25 \text{ mm} \\ 5 \text{ mm} \end{pmatrix} = s \begin{pmatrix} -2,5 \text{ mm} \\ -1,25 \text{ mm} \\ 5 \text{ mm} \end{pmatrix}$$

2.c

Im Folgenden wird die Einheit [mm] der Einfachheit wegen weggelassen. Koordinatentransformation:

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

$$\vec{x} \rightarrow \vec{x} + \begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix}$$

g_r im Weltkoordinatensystem (Richtungsvektor bleibt unverändert, da Koordinatentransformation nur translativ). Die Koordinatentransformation wird auf den Aufpunktsvektor (=Nullvektor) angewendet:

$$g_r : \vec{x} = \begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -2,5 \\ -1,25 \\ 5 \end{pmatrix}$$

Gleichsetzen der im Weltkoordinatensystem definierten Geraden g_l und g_r ergibt:

$$r \begin{pmatrix} 0 \\ -1,25 \\ 5 \end{pmatrix} = \begin{pmatrix} 100 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} -2,5 \\ -1,25 \\ 5 \end{pmatrix}$$

Hieraus ergibt sich $r = s = 40$ und somit:

$$\vec{s} = \begin{pmatrix} 0 \\ -50 \\ 200 \end{pmatrix}$$

Onlinefrage Nr. 2:

Die Koordinaten des Punktes S aus Aufgabe 2.c lauten $S(0 \text{ mm} / -50 \text{ mm} / 200 \text{ mm})$.

2.d

Dies hat zur Folge, dass sich die Geraden g_l' und g_r nicht exakt in einem Punkt schneiden, sondern windschief sind. Der optimale Schnittpunkt S' kann beispielsweise entweder als Mittelpunkt der kürzesten Verbindungsstrecke zweier windschiefer Geraden berechnet werden oder über die Lösung eines überbestimmten LGS. Letztere Variante wird im Folgenden vorgerechnet. Gegeben seien allgemein zwei Geraden g_l und g_r :

$$g_l : \vec{x} = \vec{a} + r\vec{u}$$

$$g_r : \vec{x} = \vec{b} + s\vec{v}$$

Gleichsetzen von g_l und g_r ergibt:

$$r\vec{u} - s\vec{v} = \vec{b} - \vec{a}$$

$$\Leftrightarrow \begin{pmatrix} u_1 & -v_1 \\ u_2 & -v_2 \\ u_3 & -v_3 \end{pmatrix} \begin{pmatrix} r \\ s \end{pmatrix} = \vec{b} - \vec{a}$$

Mit

$$A := \begin{pmatrix} u_1 & -v_1 \\ u_2 & -v_2 \\ u_3 & -v_3 \end{pmatrix} \quad \text{und} \quad \mathbf{x} := \begin{pmatrix} r \\ s \end{pmatrix} \quad \text{und} \quad \mathbf{b} := \vec{b} - \vec{a}$$

ist $\mathbf{Ax} = \mathbf{b}$ ein überbestimmtes LGS, dessen optimale Lösung \mathbf{x}^* im Sinne der euklidischen Norm beispielsweise durch Lösung der Normalengleichung $A^T \mathbf{Ax}^* = A^T \mathbf{b}$ zu

$$\mathbf{x}^* = (A^T A)^{-1} A^T \mathbf{b}$$

berechnet werden kann (numerisch stabilere Lösungsverfahren basieren auf der Singulärwertzerlegung oder QR-Zerlegung). Mit $(r^*, s^*) := \mathbf{x}^*$ lässt sich der Punkt mit dem geringsten Abstand zu beiden Geraden berechnen zu:

$$s' = \frac{\mathbf{a} + r^* \cdot \mathbf{u} + \mathbf{b} + s^* \cdot \mathbf{v}}{2}$$

Aufgabe 3 – Kontrastanpassung

3.a

Da sowohl die Intensität 0 als auch 255 vorhanden sind, verändert die Spreizung das Bild nicht. Es gilt also $B' = B$.

3.b

Das Histogramm $H(x)$ mit $H : \{0, \dots, 255\} \rightarrow \mathbb{Z}$ lautet:

$$H(x) = \begin{cases} 1 & x = 0 \\ 2 & x = 80 \\ 3 & x = 100 \\ 3 & x = 120 \\ 2 & x = 160 \\ 1 & x = 255 \\ 0 & \text{sonst} \end{cases}$$

Für das akkumulierte Histogramm $H_a(x)$ gilt:

$$H_a(x) = \begin{cases} 1 & 0 \leq x \leq 79 \\ 3 & 80 \leq x \leq 99 \\ 6 & 100 \leq x \leq 119 \\ 9 & 120 \leq x \leq 159 \\ 11 & 160 \leq x \leq 254 \\ 12 & x = 255 \end{cases}$$

Das 0,1-Quantil berechnet sich zu $H_q(0,1) = 80$ und das 0,9-Quantil zu $H_q(0,9) = 160$. Die Parameter der affinen Punktoperation $I'(u, v) = aI(u, v) + b$ sind also $a = \frac{255}{160-80} = 3,1875$ und $b = -\frac{255 \cdot 80}{160-80} = -255$. Damit ergibt sich die folgende neue Bildmatrix B' :

$$B' = \begin{pmatrix} 0 & 0 & 128 & 0 \\ 64 & 64 & 128 & 255 \\ 64 & 255 & 255 & 128 \end{pmatrix}$$

Onlinefrage Nr. 3:

Die Summe aller Einträge der neuen Bildmatrix beträgt 1341.

3.c

Es gilt $H_n(x) = \frac{255}{12} \cdot H_a(x)$. Anwendung von $H_n(x)$ ergibt die neue Bildmatrix B' :

$$B' = \begin{pmatrix} 64 & 64 & 191 & 21 \\ 128 & 128 & 191 & 255 \\ 128 & 234 & 234 & 191 \end{pmatrix}$$

Onlinefrage Nr. 4:

Die Summe aller Einträge der neuen Bildmatrix beträgt 1829.

Aufgabe 4 – Filter

4.a

Die Formel zur Berechnung des Gauß-Filters lautet:

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}}$$

Als Abschätzung für die Filtergröße ergibt sich: $n = \lfloor 2 \cdot 0.85 \rfloor \cdot 2 + 1 = 3$, d.h. $x, y \leq 1$. Es ergibt sich somit die folgende Filtermatrix:

$$\begin{pmatrix} 0.0552 & 0.1103 & 0.0552 \\ 0.1103 & 0.2203 & 0.1103 \\ 0.0552 & 0.1103 & 0.0552 \end{pmatrix}$$

Für den zentralen Eintrag ergibt sich im Vergleich zur gegebenen Filtermatrix der Faktor $\frac{4}{0.2203} \approx 18.16$, und es gilt $0.0552 \cdot 18.16 \approx 1$ sowie $0.1103 \cdot 18.16 \approx 2$.

Bei einem Tiefpassfilter muss die Summe der Einträge 1 betragen, und es gilt $4 + 4 \cdot 2 + 4 \cdot 1 = 16$. Es gilt zu beachten, dass in der oben berechneten Filtermatrix die Summe der Einträge einen Wert < 1 besitzt; bei unendlicher Fortsetzung würde die Summe 1 betragen.

4.b

$$B' = \begin{pmatrix} \dots & \dots \\ \dots & 20 & 20 & 28 & 43 & 50 & 43 & 28 & 20 & 20 & \dots \\ \dots & 20 & 20 & 28 & 43 & 50 & 43 & 28 & 20 & 20 & \dots \\ \dots & 20 & 20 & 28 & 43 & 50 & 43 & 28 & 20 & 20 & \dots \\ \dots & 20 & 20 & 28 & 43 & 50 & 43 & 28 & 20 & 20 & \dots \\ \dots & \dots \end{pmatrix}$$

4.c

Durch Faltung mit dem Sobel_x-Filter ergibt sich:

$$\begin{pmatrix} \dots & \dots \\ \dots & \dots \\ \dots & \dots & 32 & 92 & 88 & 0 & -88 & -92 & -32 & \dots & \dots \\ \dots & \dots & 32 & 92 & 88 & 0 & -88 & -92 & -32 & \dots & \dots \\ \dots & \dots & 32 & 92 & 88 & 0 & -88 & -92 & -32 & \dots & \dots \\ \dots & \dots \\ \dots & \dots \end{pmatrix}$$

4.d

Durch Faltung mit dem Sobel_y-Filter ergibt sich:

$$\begin{pmatrix} \dots & \dots \\ \dots & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots \\ \dots & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots \\ \dots & \dots & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & \dots \\ \dots & \dots \\ \dots & \dots \end{pmatrix}$$

Onlinefrage Nr. 5:

Der betraglich größte Eintrag der neuen Bildmatrix ist 92.

Musterlösung zu Übungsblatt 5

Bildverarbeitung - Teil 2

Aufgabe 1 – Morphologische Operatoren

1.a

```
255 255 255
255 255 255
255 255 255
```

1.a

```
0 0 0
0 0 0
0 0 0
```

Onlinefrage Nr. 1:

Die Summe der mit X markierten Elemente beträgt 2295.

Aufgabe 2 – Punkte \Rightarrow Geraden

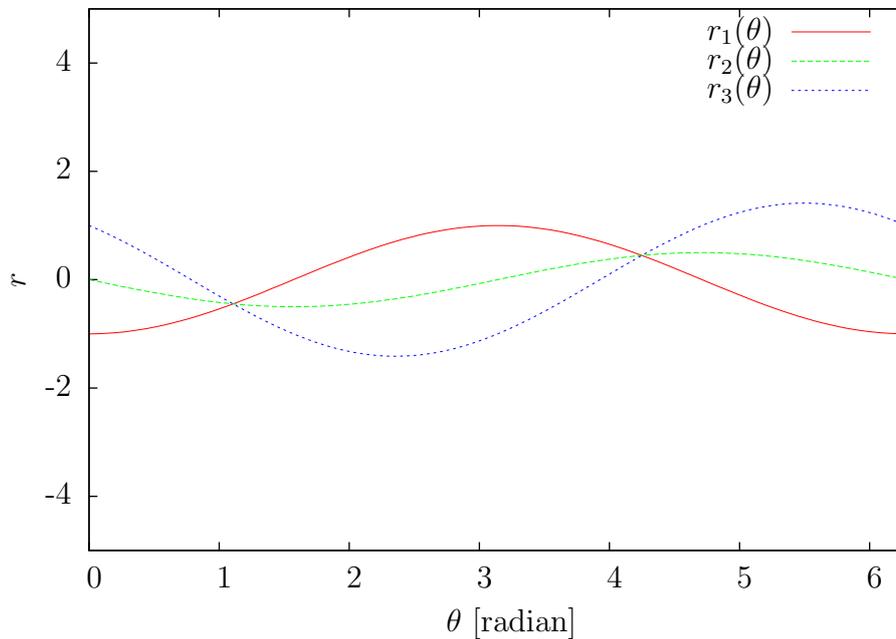
2.a

$$P_1 : r_1 = -\cos \theta$$

$$P_2 : r_2 = -0.5 \sin \theta$$

$$P_3 : r_3 = \cos \theta - \sin \theta$$

Im Hough-Diagramm lässt sich als θ -Koordinate des Schnittpunktes mit $0 \leq \theta < \pi$ in etwa der Wert 1.1 ablesen.



Als r lässt sich z.B. $r_1 = -\cos \theta \approx -\cos 1.1 \approx -0.45$ heranziehen. Ferner gilt $\sin 1.1 \approx 0.89$ und somit gilt:

$$g : \begin{pmatrix} 0.45 \\ 0.89 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -0.45$$

Ein Aufpunktvektor kann berechnet werden, indem beispielsweise $x = 0$ eingesetzt wird. Somit gilt (Richtungsvektor verläuft senkrecht zum Normalenvektor):

$$g : \mathbf{x} = \begin{pmatrix} 0 \\ -0.5 \end{pmatrix} + s \cdot \begin{pmatrix} 0.89 \\ -0.45 \end{pmatrix}$$

2.b

Die Bestimmung der Regressionsgerade ist das geeignetere Verfahren, wenn schon klar ist, welche Pixel eine Gerade bilden sollen; dann ist die Regressionsgerade die optimale Gerade im Sinne der Summe der Fehlerquadrate. Die Regressionsgerade kann jedoch keine Segmentierung vornehmen. Dahingegen lassen sich mit der Hough-Transformation in beliebigen Bildern geradlinige Strukturen erkennen; die Punkte dazu müssen dazu jedoch verhältnismässig exakt auf einer Linie liegen.

Onlinefrage Nr. 2:

$\theta \approx 1.1$.

Aufgabe 3 – Korrelation

3.a

$$\begin{aligned}SAD(A, B_1) &= 4 \cdot 45 = 180 \\SAD(A, B_2) &= 90 + 45 = 135\end{aligned}$$

Als Korrespondenz wird folglich B_2 bestimmt.

3.b

$$\begin{aligned}SSD(A, B_1) &= 4 \cdot 45^2 = 8100 \\SSD(A, B_2) &= 90^2 + 45^2 = 10125\end{aligned}$$

Als Korrespondenz wird folglich B_1 bestimmt.

3.c

$$\begin{aligned}\bar{A} &= 20 \\ \bar{B}_1 &= 40 \\ \bar{B}_2 &= 25 \\ ZNCC(A, B_1) &= \frac{4 \cdot (25 \cdot 50) + 5 \cdot ((-20) \cdot (-40))}{\sqrt{(4 \cdot 25^2 + 5 \cdot (-20)^2) \cdot (4 \cdot 50^2 + 5 \cdot 40^2)}} = \frac{5000 + 4000}{\sqrt{4500 \cdot 18000}} = \frac{9000}{9000} = 1 \\ ZNCC(A, B_2) &= \frac{3 \cdot (25 \cdot 20) + (25 \cdot (-25)) + ((-20) \cdot 65) + 4 \cdot ((-20) \cdot (-25))}{\sqrt{(4 \cdot 25^2 + 5 \cdot (-20)^2) \cdot (3 \cdot 20^2 + 65^2 + 5 \cdot (-25)^2)}} \\ &= \frac{1500 - 625 - 1300 + 2000}{\sqrt{4500 \cdot 8550}} \approx \frac{1575}{6203} \approx 0.254\end{aligned}$$

Als Korrespondenz wird folglich B_1 bestimmt.

Onlinefrage Nr. 3:

Das Ergebnis der Korrelation in Teilaufgabe 3.c für die berechnete Korrespondenz B_1 beträgt 1.

Aufgabe 4 – 3D-Transformationen

4.a

$$\begin{aligned}R_{X'Z'Y'}(\alpha, \beta, \gamma) &= R_X(\alpha) \cdot R_Z(\beta) \cdot R_Y(\gamma) \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c\alpha & -s\alpha \\ 0 & s\alpha & c\alpha \end{pmatrix} \begin{pmatrix} c\beta & -s\beta & 0 \\ s\beta & c\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c\gamma & 0 & s\gamma \\ 0 & 1 & 0 \\ -s\gamma & 0 & c\gamma \end{pmatrix}\end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} c\beta & -s\beta & 0 \\ c\alpha s\beta & c\alpha c\beta & -s\alpha \\ s\alpha s\beta & s\alpha c\beta & c\alpha \end{pmatrix} \begin{pmatrix} c\gamma & 0 & s\gamma \\ 0 & 1 & 0 \\ -s\gamma & 0 & c\gamma \end{pmatrix} \\
&= \begin{pmatrix} c\beta c\gamma & -s\beta & c\beta s\gamma \\ c\alpha s\beta c\gamma + s\alpha s\gamma & c\alpha c\beta & c\alpha s\beta s\gamma - s\alpha c\gamma \\ s\alpha s\beta c\gamma - c\alpha s\gamma & s\alpha c\beta & s\alpha s\beta s\gamma + c\alpha c\gamma \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
R_{X'Z'Y'}(\alpha, \beta, \gamma)^{-1} &= R_{X'Z'Y'}(\alpha, \beta, \gamma)^T \\
&= \begin{pmatrix} c\beta c\gamma & c\alpha s\beta c\gamma + s\alpha s\gamma & s\alpha s\beta c\gamma - c\alpha s\gamma \\ -s\beta & c\alpha c\beta & s\alpha c\beta \\ c\beta s\gamma & c\alpha s\beta s\gamma - s\alpha c\gamma & s\alpha s\beta s\gamma + c\alpha c\gamma \end{pmatrix}
\end{aligned}$$

4.b

$$R_{YZX}(\gamma, \beta, \alpha) = R_X(\alpha) \cdot R_Z(\beta) \cdot R_Y(\gamma) = R_{X'Z'Y'}(\alpha, \beta, \gamma) \quad (\text{s. 4.a})$$

Aufgabe 5 – Kameramodell

5.a

Es gilt:

$$\begin{aligned}
\mathbf{x}_1 &= R_1 \mathbf{x}_w + \mathbf{t}_1 \\
\mathbf{x}_2 &= R_2 \mathbf{x}_w + \mathbf{t}_2
\end{aligned}$$

Dementsprechend gilt für die inverse Transformation:

$$\begin{aligned}
\mathbf{x}_w &= R_1^T \mathbf{x}_1 - R_1^T \mathbf{t}_1 \\
\mathbf{x}_w &= R_2^T \mathbf{x}_2 - R_2^T \mathbf{t}_2
\end{aligned}$$

Die Projektionszentren der Kameras haben jeweils in ihrem eigenen Koordinatensystem den Nullvektor als Ortsvektor. Somit gilt:

$$\begin{aligned}
\mathbf{c}_{1,w} &= R_1^T \mathbf{0} - R_1^T \mathbf{t}_1 = -R_1^T \mathbf{t}_1 \\
\mathbf{c}_{2,w} &= R_2^T \mathbf{0} - R_2^T \mathbf{t}_2 = -R_2^T \mathbf{t}_2
\end{aligned}$$

Und schließlich:

$$\overrightarrow{C_1 C_{2w}} = \mathbf{c}_{2,w} - \mathbf{c}_{1,w} = R_1^T \mathbf{t}_1 - R_2^T \mathbf{t}_2$$

5.b

Wenn das neue Weltkoordinatensystem identisch mit dem Kamerakoordinatensystem der Kamera K_1 sein soll, dann gilt:

$$\begin{aligned}
R'_1 &= I \\
\mathbf{t}'_1 &= \mathbf{0}
\end{aligned}$$

Gesucht ist nun für R'_2, \mathbf{t}'_2 die Koordinatentransformation vom Kamerakoordinatensystem der Kamera K_1 in das Kamerakoordinatensystem der Kamera K_2 . Diese lässt sich über das alte Weltkoordinatensystem wie folgt berechnen. Es gilt (s. 5.a):

$$\begin{aligned}\mathbf{x}_w &= R_1^T \mathbf{x}_1 - R_1^T \mathbf{t}_1 \\ \mathbf{x}_2 &= R_2 \mathbf{x}_w + \mathbf{t}_2\end{aligned}$$

Daraus folgt:

$$\mathbf{x}_2 = R_2 R_1^T \mathbf{x}_1 - R_2 R_1^T \mathbf{t}_1 + \mathbf{t}_2$$

Und somit gilt:

$$\begin{aligned}R'_2 &= R_2 R_1^T \\ \mathbf{t}'_2 &= \mathbf{t}_2 - R_2 R_1^T \mathbf{t}_1\end{aligned}$$

Für die Ortsvektoren der Projektionszentren im neuen Weltkoordinatensystem gilt:

$$\begin{aligned}\mathbf{c}_{1,w'} &= \mathbf{0} \\ \mathbf{c}_{2,w'} &= R_2'^T \mathbf{0} - R_2'^T \mathbf{t}'_2 = -R_2'^T \mathbf{t}'_2 \\ &= -(R_2 R_1^T)^T (\mathbf{t}_2 - R_2 R_1^T \mathbf{t}_1) \\ &= R_1 R_2^T (R_2 R_1^T \mathbf{t}_1 - \mathbf{t}_2) \\ &= \mathbf{t}_1 - R_1 R_2^T \mathbf{t}_2\end{aligned}$$

Und schließlich:

$$\overrightarrow{C_1 C_{2w'}} = \mathbf{c}_{2,w'} - \mathbf{c}_{1,w'} = \mathbf{t}_1 - R_1 R_2^T \mathbf{t}_2$$

5.c

$$T = \begin{pmatrix} R & \mathbf{t} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$T^{-1} = \begin{pmatrix} R^T & -R^T \mathbf{t} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Onlinefrage Nr. 4:

Das Ergebnis lautet $R_1^T \mathbf{t}_1 - R_2^T \mathbf{t}_2$.

Aufgabe 6 – Epipolargeometrie

6.a

Wie auf dem alten Blatt berechnet, ist der Umrechnungsfaktor von [mm] nach [Pixel] $100 \frac{\text{Pixel}}{\text{mm}}$. Die Brennweite ist gegeben mit $f = 5$ mm. Daraus folgt:

$$f_x = f_y = 5 \text{ mm} \cdot 100 \frac{\text{Pixel}}{\text{mm}} = 500 \text{ Pixel.}$$

Somit gilt für die Kalibriermatrizen K_1, K_2 (Einheit ist [Pixel]):

$$K_1 = K_2 = \begin{pmatrix} 500 & 0 & 320 \\ 0 & 500 & 240 \\ 0 & 0 & 1 \end{pmatrix}$$

Für die extrinsischen Parameter gilt (Einheit für die translativen Parameter ist [mm]):

$$\begin{aligned} R_1 &= I, \mathbf{t}_1 = \mathbf{0} \\ R_2 &= I, \mathbf{t}_2 = (-100 \ 0 \ 0)^T \end{aligned}$$

Die Projektionsmatrizen lauten dann:

$$\begin{aligned} P_1 &= \begin{pmatrix} 500 & 0 & 320 & 0 \\ 0 & 500 & 240 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\ P_2 &= \begin{pmatrix} 500 & 0 & 320 & -50000 \\ 0 & 500 & 240 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{aligned}$$

6.b

$$E = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix} \Rightarrow \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 100 \\ 0 & -100 & 0 \end{pmatrix}$$

Um die Fundamentalmatrix F auf der Basis der Essentialmatrix E zu berechnen, muss zunächst $K_1^{-1} = K_2^{-1}$ berechnet werden:

$$\begin{aligned} &\left(\begin{array}{ccc|ccc} 500 & 0 & 320 & 1 & 0 & 0 \\ 0 & 500 & 240 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \\ \rightsquigarrow &\left(\begin{array}{ccc|ccc} 500 & 0 & 0 & 1 & 0 & -320 \\ 0 & 500 & 0 & 0 & 1 & -240 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \\ \rightsquigarrow &\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0.002 & 0 & -0.64 \\ 0 & 1 & 0 & 0 & 0.002 & -0.48 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \\ \Rightarrow K_1^{-1} = K_2^{-1} &= \begin{pmatrix} 0.002 & 0 & -0.64 \\ 0 & 0.002 & -0.48 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Die Fundamentalmatrix lässt sich nun wie folgt berechnen:

$$\begin{aligned} F &= K_2^{-T} E K_1^{-1} \\ &= \begin{pmatrix} 0.002 & 0 & 0 \\ 0 & 0.002 & 0 \\ -0.64 & -0.48 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 100 \\ 0 & -100 & 0 \end{pmatrix} \begin{pmatrix} 0.002 & 0 & -0.64 \\ 0 & 0.002 & -0.48 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0.2 \\ 0 & -100 & -48 \end{pmatrix} \begin{pmatrix} 0.002 & 0 & -0.64 \\ 0 & 0.002 & -0.48 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0.2 \\ 0 & -0.2 & 0 \end{pmatrix}$$

Für die Epipole gilt:

$$\mathbf{e}_1 = -K_1 R_2^T \mathbf{t}_2 = - \begin{pmatrix} 500 & 0 & 320 \\ 0 & 500 & 240 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -100 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 50000 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{e}_2 = K_2 \mathbf{t}_2 = \begin{pmatrix} 500 & 0 & 320 \\ 0 & 500 & 240 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -100 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -50000 \\ 0 \\ 0 \end{pmatrix}$$

Division durch die w -Komponente des Ergebnisses bedeutet eine Division durch Null. D.h. die Epipole liegen im Unendlichen auf der u -Achse, weshalb alle Epipolarlinien horizontal verlaufen.

6.c

$$\mathbf{l}_2 = F \begin{pmatrix} 200 \\ 200 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0.2 \\ 0 & -0.2 & 0 \end{pmatrix} \begin{pmatrix} 200 \\ 200 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.2 \\ -40 \end{pmatrix}$$

Onlinefrage Nr. 5:

Das Ergebnis lautet $(0 \ 0.2 \ -40)^T$.

Musterlösung zu Übungsblatt 6

Wissensrepräsentation, Bahnplanung und Robotik

Aufgabe 6.1 Resolution und DPLL

6.1.a

Jeder Satz läßt sich, soweit er erfüllbar ist, durch eine Wahrheitstabelle darstellen. Daraus gewinnt man eine KF wie folgt:

Alle Zeilen deren Gesamtwahrheitswert FALSE ist, werden konjunktiv verknüpft. Die Konjunktionen selbst bestehen dabei aus den disjunktiv verknüpften und negiert oder nicht negierten Variablen aus der Tabelle.

Für die Sätze α , β und γ ergibt sich folgende Wertetabelle:

α	β	γ	A	B	C
1	0	1	0	0	0
1	1	0	0	0	1
1	0	1	0	1	0
1	1	0	0	1	1
0	0	1	1	0	0
0	0	0	1	0	1
0	1	1	1	1	0
1	1	1	1	1	1

Daraus ergibt sich die KF für α , β und γ :

$$\alpha_{KF} = (\neg A \vee B \vee C) \wedge (\neg A \vee B \vee \neg C) \wedge (\neg A \vee \neg B \vee C)$$

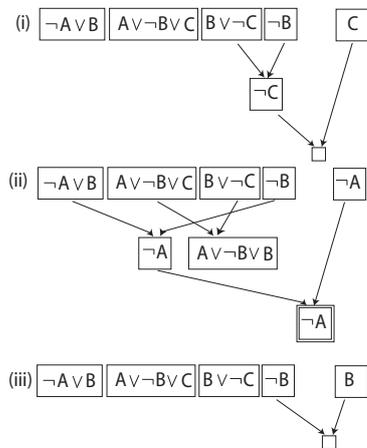
$$\beta_{KF} = (A \vee B \vee C) \wedge (A \vee \neg B \vee C) \wedge (\neg A \vee B \vee C) \wedge (\neg A \vee B \vee \neg C)$$

$$\gamma_{KF} = (A \vee B \vee \neg C) \wedge (A \vee \neg B \vee \neg C) \wedge (\neg A \vee B \vee \neg C)$$

6.1.b

Onlinefrage Nr. 1:

A kann nicht abgeleitet werden.



6.1.c

Symbol A kommt in einer Einheitsklausel vor. Setze $A=TRUE$. Dies ergibt

$$\begin{array}{l}
 B \vee \neg C \\
 \neg C \vee \neg B \vee E \\
 \neg D \vee \neg B \vee C \\
 \neg E \vee \neg D \vee B \\
 B \vee E \vee C
 \end{array}$$

Keine Einheitsklauseln. Das Symbol D kommt nur negiert vor, ist also ein reines Symbol. Wähle deshalb $D=FALSE$. Dies ergibt:

$$\begin{array}{l}
 B \vee \neg C \\
 \neg C \vee \neg B \vee E \\
 B \vee E \vee C
 \end{array}$$

Keine Einheitsklauseln. Das Symbol E kommt nur nicht negiert vor, ist also ein reines Symbol. Wähle deshalb $E=TRUE$. Dies ergibt:

$$B \vee \neg C$$

Das Symbol B (oder C) kommt nur (nicht) negiert vor, ist also ein reines Symbol. Wähle deshalb $B=TRUE$. Dadurch sind alle Klauseln erfüllt.

Ein mögliches Modell lautet somit

$$M_1 = \{\{A = TRUE\}, \{B = TRUE\}, \{C = TRUE \vee FALSE\}, \{D = FALSE\}, \{E = TRUE\}\}.$$

Alternativ kann $C=FALSE$ gewählt werden. Auch hierdurch sind alle Klauseln erfüllt. Ein weiteres mögliches Modell lautet somit

$$M_2 = \{\{A = TRUE\}, \{B = TRUE \vee FALSE\}, \{C = FALSE\}, \{D = FALSE\}, \{E = TRUE\}\}.$$

Da ein Modell für die Klauselmenge gefunden wurde, ist sie also erfüllbar.

Bemerkung: Ein weiteres Modell, welches die Klauselmenge erfüllt und dabei keine Aussagen über E macht, lautet

$$M_3 = \{\{A = TRUE\}, \{B = TRUE\}, \{C = FALSE\}\{D = FALSE\}, \{E = TRUE \vee FALSE\}\}.$$

Dieses Modell wird jedoch durch den DPLL-Algorithmus nicht gefunden. Beachte: Der DPLL-Algorithmus prüft nur die (Un-)Erfüllbarkeit einer Klauselmenge. Er findet aber nicht alle Modelle, die diese Klauselmenge erfüllen.

Onlinefrage Nr. 2:

Es existiert sowohl für $C = TRUE$ als auch $C = FALSE$ ein Modell, das S erfüllt, und das vom DPLL-Algorithmus gefunden wird.

Aufgabe 6.2 - STRIPS & ADL

6.2.a

Action(Load(c, p, a),

Vorbed : $At(c, a) \wedge At(p, a) \wedge Fracht(c) \wedge Flugzeug(p) \wedge Flughafen(a)$

Effekt : $\neg At(c, a) \wedge In(c, p)$)

Action(Unload(c, p, a),

Vorbed : $In(c, p) \wedge At(p, a) \wedge Fracht(c) \wedge Flugzeug(p) \wedge Flughafen(a)$

Effekt : $At(c, a) \wedge \neg In(c, p)$)

Action(Fly($p, from, to$),

Vorbed : $At(p, from) \wedge Flugzeug(p) \wedge Flughafen(from) \wedge Flughafen(to)$

Effekt : $\neg At(p, from) \wedge At(p, to)$)

6.2.b

Init($At(C_1, FRA) \wedge At(C_2, JFK) \wedge At(P_1, FRA) \wedge Fracht(C_1) \wedge Fracht(C_2) \wedge Flugzeug(P_1) \wedge Flughafen(FRA) \wedge Flughafen(JFK)$)

Ziel($At(C_1, JFK) \wedge At(C_2, FRA)$)

Diese Aktionsfolge ist eine Lösung des Planungsproblems:

[Load(C_1, P_1, FRA), Fly(P_1, FRA, JFK), Unload(C_1, P_1, JFK),
Load(C_2, P_1, JFK), Fly(P_1, JFK, FRA), Unload(C_2, P_1, FRA)]

6.2.c

Init($At(C_1, FRA) \wedge At(C_2, JFK) \wedge At(P_1, FRA) \wedge Fracht(C_1) \wedge Fracht(C_2) \wedge Flugzeug(P_1) \wedge Flughafen(FRA) \wedge Flughafen(JFK)$)

Ziel($\exists x At(C_1, x) \wedge At(C_2, x)$)

Diese Aktionsfolge ist eine Lösung des Planungsproblems:

[Load(C_1, P_1, FRA), Fly(P_1, FRA, JFK), Unload(C_1, P_1, JFK)]

Aufgabe 6.3 - Bahnplanung

6.3.a

Die entsprechenden Räume sind in Abb. 2 jeweils in grau zu sehen. Der Konfigurationsraum besteht aus allen Punkten die für den Roboter in der Welt erreichbar sind. In diesem Falle also aus allen Punkten innerhalb der gegebenen quadratischen Welt (Abb. 2 a). Der

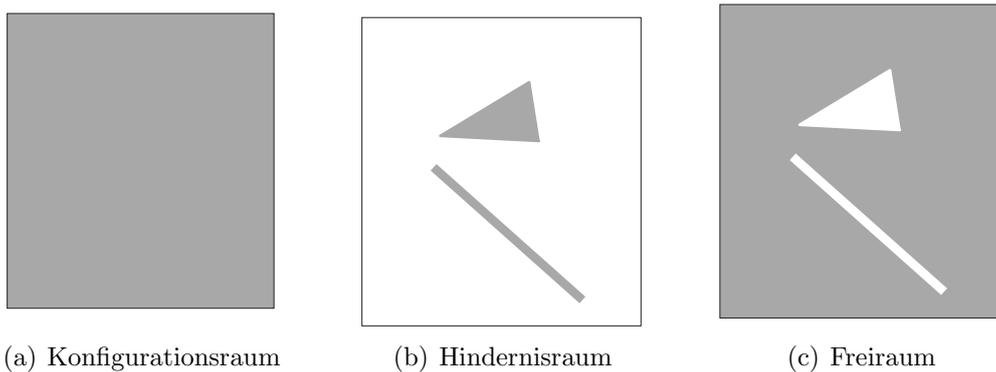


Abbildung 1: Räume zur Bahnplanung

Hindernisraum besteht aus den Konfigurationen, bei denen eine Berührung oder eine Durchdringung des Roboters mit dem Hindernis vorläge (Abb. 2 b). Der Freiraum besteht aus dem Konfigurationsraum ohne den Hindernisraum (Abb. 2 c). Der angegebene Raum ist jeweils in grau dargestellt. Zur Darstellung des Roboters als beweglichen Punkt können alle Hindernisse um den Roboterradius $\frac{d}{2}$ erweitert werden und auf diesem Modell Frei- und Hindernisraum berechnet werden.

6.3.b

Im Sichtgraph sind die Verbindungen aller sichtbaren Eckpunkte der Hindernisse (und des Start- und Zielpunktes) als mögliche Pfade dargestellt (Abb. 2). Auf den hierdurch ermittelten Pfaden kann sich der mobile Roboter kollisionsfrei bewegen.

Onlinefrage Nr. 3:

Der Graph besitzt 19 Kanten.

6.3.c

Planer 1 verwendet den Greedy-Algorithmus. Er ermittelt einen Weg über Start-G-D-E-Ziel mit einer Gesamtlänge von 236 (siehe Abb.3). Planer 2, der den A*-Algorithmus anwendet, ermittelt einen Weg über Start-B-E-Ziel mit einer Gesamtlänge von 115 (siehe Abb. 4). Mit * markierte Knoten werden nicht weiterverfolgt, da sie bereits weiter oben im Baum durchlaufen wurden.

Onlinefrage Nr. 4:

Die Wegstrecke beträgt 115 Einheiten.

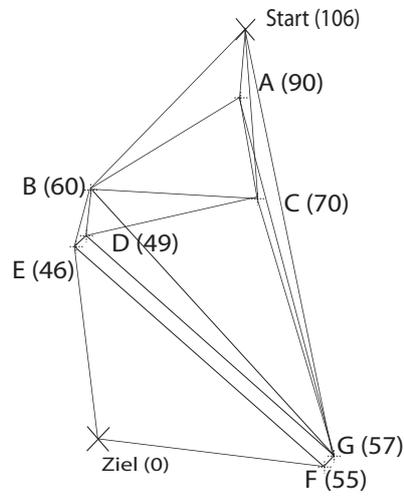


Abbildung 2: Sichtgraph

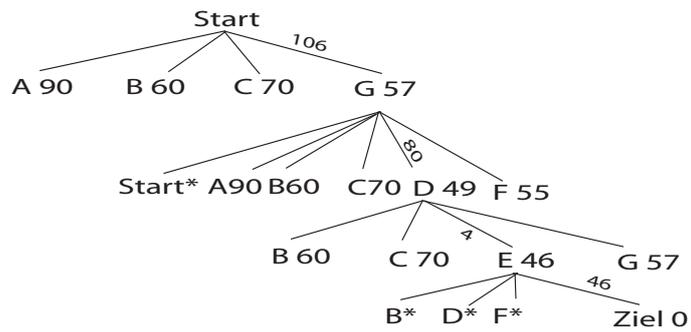


Abbildung 3: Suchbaum Planer 1

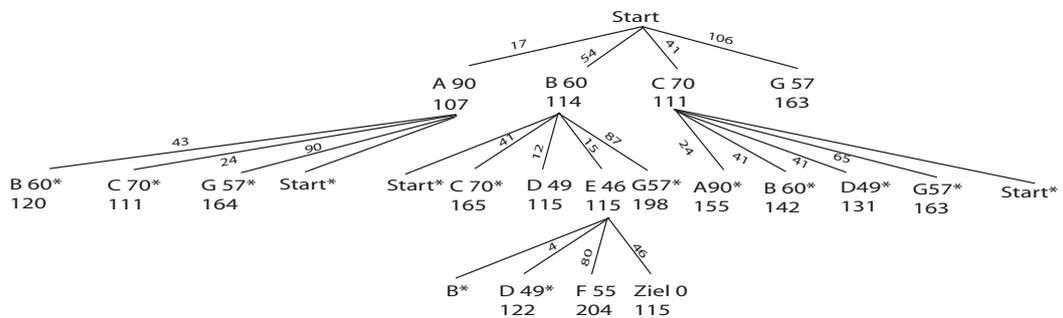


Abbildung 4: Suchbaum Planer 2

6.3.d

Siehe Abbildung 5.

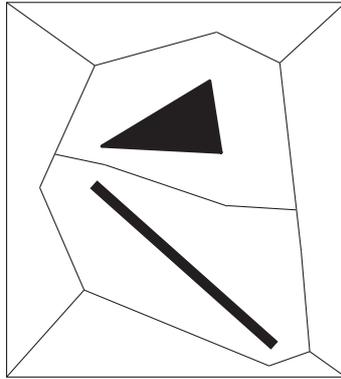


Abbildung 5: Voronoi-Diagramm

Aufgabe 6.4 - Quaternionen

6.4.a

Mit $\bar{q}_1 = (2, (-4, 3, 0))$ und $|q_1|^2 = 4 + 16 + 9 = 29$ errechnet sich das inverse Quaternion q_1^{-1} zu:

$$q_1^{-1} = \frac{\bar{q}_1}{|q_1|^2} = \left(\frac{2}{29}, \left(-\frac{4}{29}, \frac{3}{29}, 0\right) \right)$$

6.4.b

Mit \vec{x} als Quaternion $\vec{x} = (0, (0, 0, 3))$ und $\bar{q}_2 = \left(\frac{\sqrt{2}}{2}, \left(-\frac{\sqrt{2}}{2}, 0, 0\right)\right)$ berechnet sich der rotierte Punkt \vec{x}' wie folgt:

$$\begin{aligned} \vec{x}' &= q_2 \vec{x} \bar{q}_2 \\ &= \left(\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i\right) \cdot 3k \cdot \left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i\right) \\ &= \left(\frac{3\sqrt{2}}{2}k - \frac{3\sqrt{2}}{2}j\right) \cdot \left(\frac{\sqrt{2}}{2} - \frac{\sqrt{2}}{2}i\right) \\ &= \frac{3}{2}k - \frac{3}{2}j - \frac{3}{2}j - \frac{3}{2}k \\ &= -3j \\ \Rightarrow \vec{x}' &= (0, -3, 0) \end{aligned}$$

6.4.c

Mit Quaternionen lassen sich nur Rotationen darstellen, keine Translationen. Dies ist ein großer Nachteil von Quaternionen.

Onlinefrage Nr. 5: Nur Rotationen.